Chapter 2
Simple Comparative Experiments
Solutions

2.1. Computer output for a random sample of data is shown below. Some of the quantities are missing. Compute the values of the missing quantities.

Variable	N	Mean	SE Mean	Std. Dev.	Variance	Minimum	Maximum
Y	9	19.96	?	3.12	?	15.94	27.16

SE Mean = 1.04 Variance = 9.73

2.2. Computer output for a random sample of data is shown below. Some of the quantities are missing. Compute the values of the missing quantities.

Variable	N	Mean	SE Mean	Std. Dev.	Sum
Y	16	?	0.159	?	399.851

Mean = 24.991 Std. Dev. = 0.636

2.3. Suppose that we are testing H_0: $\mu = \mu_0$ versus H_1: $\mu \neq \mu_0$. Calculate the *P*-value for the following observed values of the test statistic:

(a) $Z_0 = 2.25$ *P*-value = 0.02445

(b) $Z_0 = 1.55$ *P*-value = 0.12114

(c) $Z_0 = 2.10$ *P*-value = 0.03573

(d) $Z_0 = 1.95$ *P*-value = 0.05118

(e) $Z_0 = -0.10$ *P*-value = 0.92034

2.4. Suppose that we are testing H_0: $\mu = \mu_0$ versus H_1: $\mu > \mu_0$. Calculate the *P*-value for the following observed values of the test statistic:

(a) $Z_0 = 2.45$ *P*-value = 0.00714

(b) $Z_0 = -1.53$ *P*-value = 0.93699

(c) $Z_0 = 2.15$ *P*-value = 0.01578

(d) $Z_0 = 1.95$ *P*-value = 0.02559

(e) $Z_0 = -0.25$ *P*-value = 0.59871

2.5. Consider the computer output shown below.

One-Sample Z

Test of mu = 30 vs not = 30

The assumed standard deviation = 1.2

N	Mean	SE Mean	95% CI	Z	P
16	31.2000	0.3000	(30.6120, 31.7880)	?	?

(a) Fill in the missing values in the output. What conclusion would you draw?

Z = 4 P = 0.00006; therefore, the mean is not equal to 30.

(b) Is this a one-sided or two-sided test?

Two-sided.

(c) Use the output and the normal table to find a 99 percent CI on the mean.

CI = 30.42725, 31.97275

(d) What is the *P*-value if the alternative hypothesis is H_1: $\mu > 30$

P-value = 0.00003

2.6. Suppose that we are testing H_0: $\mu_1 = \mu_2$ versus H_1: $\mu_1 = \mu_2$ with a sample size of $n_1 = n_2 = 12$. Both sample variances are unknown but assumed equal. Find bounds on the *P*-value for the following observed values of the test statistic:

(a) $t_0 = 2.30$ Table *P*-value = 0.02, 0.05 Computer *P*-value = 0.0313

(b) $t_0 = 3.41$ Table *P*-value = 0.002, 0.005 Computer *P*-value = 0.0025

(c) $t_0 = 1.95$ Table *P*-value = 0.1, 0.05 Computer *P*-value = 0.0640

(d) $t_0 = -2.45$ Table *P*-value = 0.05, 0.02 Computer *P*-value = 0.0227

Note that the degrees of freedom is (12 +12) – 2 = 22. This is a two-sided test

2.7. Suppose that we are testing H_0: $\mu_1 = \mu_2$ versus H_1: $\mu_1 > \mu_2$ with a sample size of $n_1 = n_2 = 10$. Both sample variances are unknown but assumed equal. Find bounds on the *P*-value for the following observed values of the test statistic:

(a) $t_0 = 2.31$ Table *P*-value = 0.01, 0.025 Computer *P*-value = 0.01648

(b) $t_0 = 3.60$ Table *P*-value = 0.001, 0.0005 Computer *P*-value = 0.00102

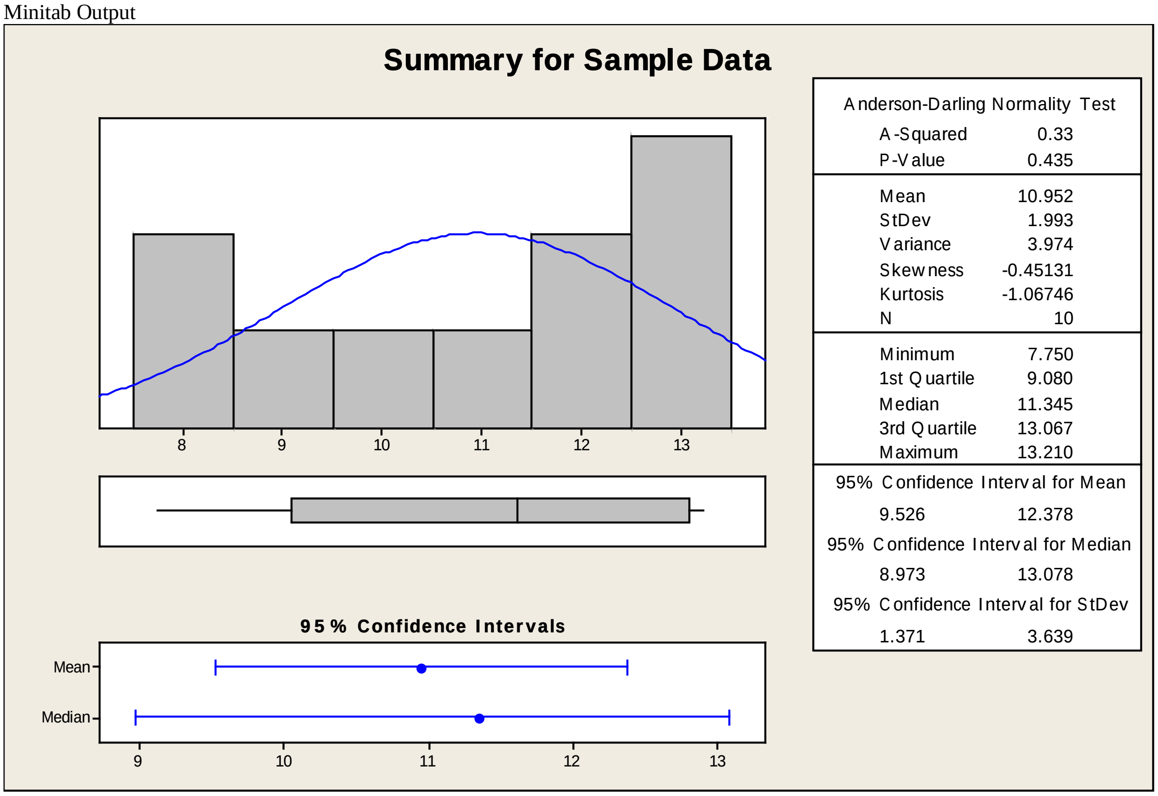
(c) $t_0 = 1.95$ Table *P*-value = 0.05, 0.025 Computer *P*-value = 0.03346

(d) $t_0 = 2.19$ Table P-value = 0.01, 0.025 Computer P-value = 0.02097

Note that the degrees of freedom is (10 +10) – 2 = 18. This is a one-sided test.

2.8. Consider the following sample data: 9.37, 13.04, 11.69, 8.21, 11.18, 10.41, 13.15, 11.51, 13.21, and 7.75. Is it reasonable to assume that this data is from a normal distribution? Is there evidence to support a claim that the mean of the population is 10?

Minitab Output



According to the output, the Anderson-Darling Normality Test has a P-Value of 0.435. The data can be considered normal. The 95% confidence interval on the mean is (9.526,12.378). This confidence interval contains 10, therefore there is evidence that the population mean is 10.

2.9. A computer program has produced the following output for the hypothesis testing problem:

Difference in sample means: 2.35
Degrees of freedom: 18
Standard error of the difference in the sample means: ?
Test statistic: $t_o = 2.01$
P-Value = 0.0298

(a) What is the missing value for the standard error?

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{2.35}{StdError} = 2.01$$

$$StdError = 2.35/2.01 = 1.169$$

(b) Is this a two-sided or one-sided test? One-sided test for a t_0 = 2.01 is a *P*-value of 0.0298.

(c) If α=0.05, what are your conclusions? Reject the null hypothesis and conclude that there is a difference in the two samples.

(d) Find a 90% two-sided CI on the difference in the means.

$$\bar{y}_1 - \bar{y}_2 - t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_1 \le \bar{y}_1 - \bar{y}_2 + t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$\bar{y}_1 - \bar{y}_2 - t_{0.05,18}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} \le \mu_1 - \mu_1 \le \bar{y}_1 - \bar{y}_2 + t_{0.05,18}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

$$2.35 - 1.734(1.169) \le \mu_1 - \mu_1 \le 2.35 + 1.734(1.169)$$

$$0.323 \le \mu_1 - \mu_1 \le 4.377$$

2.10. A computer program has produced the following output for the hypothesis testing problem:

Difference in sample means: 11.5
Degrees of freedom: 24
Standard error of the difference in the sample means: ?
Test statistic: t_o = -1.88
P-Value = 0.0723

(a) What is the missing value for the standard error?

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{-11.5}{StdError} = -1.88$$

$$StdError = -11.5/-1.88 = 6.12$$

(b) Is this a two-sided or one-sided test? Two-sided test for a t_0 = -1.88 is a *P*-value of 0.0723.

(c) If α=0.05, what are your conclusions? Accept the null hypothesis, there is no difference in the means.

(d) Find a 90% two-sided CI on the difference in the means.

$$\bar{y}_1-\bar{y}_2-t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le \mu_1-\mu_1\le \bar{y}_1-\bar{y}_2+t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$
$$\bar{y}_1-\bar{y}_2-t_{0.05,24}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\le \mu_1-\mu_1\le \bar{y}_1-\bar{y}_2+t_{0.05,24}S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$
$$-11.5-1.711(6.12)\le \mu_1-\mu_1\le -11.5+1.711(6.12)$$
$$-21.97\le \mu_1-\mu_1\le -1.03$$

2.11. Suppose that we are testing H_0: $\mu = \mu_0$ versus H_1: $\mu > \mu_0$ with a sample size of n = 15. Calculate bounds on the P-value for the following observed values of the test statistic:

(a)	t_0 = 2.35	Table P-value = 0.01, 0.025	Computer P-value = 0.01698
(b)	t_0 = 3.55	Table P-value = 0.001, 0.0025	Computer P-value = 0.00160
(c)	t_0 = 2.00	Table P-value = 0.025, 0.005	Computer P-value = 0.03264
(d)	t_0 = 1.55	Table P-value = 0.05, 0.10	Computer P-value = 0.07172

The degrees of freedom are 15 – 1 = 14. This is a one-sided test.

2.12. Suppose that we are testing H_0: $\mu = \mu_0$ versus H_1: $\mu \neq \mu_0$ with a sample size of n = 10. Calculate bounds on the P-value for the following observed values of the test statistic:

(a)	t_0 = 2.48	Table P-value = 0.02, 0.05	Computer P-value = 0.03499
(b)	t_0 = -3.95	Table P-value = 0.002, 0.005	Computer P-value = 0.00335
(c)	t_0 = 2.69	Table P-value = 0.02, 0.05	Computer P-value = 0.02480
(d)	t_0 = 1.88	Table P-value = 0.05, 0.10	Computer P-value = 0.09281
(e)	t_0 = -1.25	Table P-value = 0.20, 0.50	Computer P-value = 0.24282

2.13. Consider the computer output shown below.

One-Sample T: Y

Test of mu = 91 vs. not = 91

Variable	N	Mean	Std. Dev.	SE Mean	95% CI	T	P
Y	25	92.5805	?	0.4675	(91.6160, ?)	3.38	0.002

(a) Fill in the missing values in the output. Can the null hypothesis be rejected at the 0.05 level? Why?

Std. Dev. = 2.3365 UCI = 93.5450
Yes, the null hypothesis can be rejected at the 0.05 level because the P-value is much lower at 0.002.

(b) Is this a one-sided or two-sided test?

Two-sided.

(c) If the hypothesis had been H_0: $\mu = 90$ versus H_1: $\mu \neq 90$ would you reject the null hypothesis at the 0.05 level?

Yes.

(d) Use the output and the t table to find a 99 percent two-sided CI on the mean.

CI = 91.2735, 93.8875

(e) What is the P-value if the alternative hypothesis is H_1: $\mu > 91$?

P-value = 0.001.

2.14. Consider the computer output shown below.

One-Sample T: Y

Test of mu = 25 vs > 25

Variable	N	Mean	Std. Dev.	SE Mean	95% Lower Bound	T	P
Y	12	25.6818	?	0.3360	?	?	0.034

(a) How many degrees of freedom are there on the t-test statistic?

(N-1) = (12 – 1) = 11

(b) Fill in the missing information.

Std. Dev. = 1.1639 95% Lower Bound = 2.0292

2.15. Consider the computer output shown below.

Two-Sample T-Test and CI: Y1, Y2

Two-sample T for Y1 vs Y2

	N	Mean	Std. Dev.	SE Mean
Y1	20	50.19	1.71	0.38
Y2	20	52.52	2.48	0.55

Difference = mu (X1) – mu (X2)

Estimate for difference: -2.33341

95% CI for difference: (-3.69547, -0.97135)

T-Test of difference = 0 (vs not =) : T-Value = -3.47

P-Value = 0.01 DF = 38

Both use Pooled Std. Dev. = 2.1277

(a) Can the null hypothesis be rejected at the 0.05 level? Why?

Yes, the *P*-Value of 0.001 is much less than 0.05.

(b) Is this a one-sided or two-sided test?

Two-sided.

(c) If the hypothesis had been H_0: $\mu_1 - \mu_2 = 2$ versus H_1: $\mu_1 - \mu_2 \neq 2$ would you reject the null hypothesis at the 0.05 level?

Yes.

(d) If the hypothesis had been H_0: $\mu_1 - \mu_2 = 2$ versus H_1: $\mu_1 - \mu_2 < 2$ would you reject the null hypothesis at the 0.05 level? Can you answer this question without doing any additional calculations? Why?

Yes, no additional calculations are required because the test is naturally becoming more significant with the change from -2.33341 to -4.33341.

(e) Use the output and the *t* table to find a 95 percent upper confidence bound on the difference in means?

95% upper confidence bound = -1.21.

(f) What is the *P*-value if the alternative hypotheses are H_0: $\mu_1 - \mu_2 = 2$ versus H_1: $\mu_1 - \mu_2 \neq 2$?

P-value = 1.4E-07.

2.16. The breaking strength of a fiber is required to be at least 150 psi. Past experience has indicated that the standard deviation of breaking strength is $\sigma = 3$ psi. A random sample of four specimens is tested. The results are y_1=145, y_2=153, y_3=150 and y_4=147.

(a) State the hypotheses that you think should be tested in this experiment.

$$H_0: \mu = 150 \qquad H_1: \mu > 150$$

(b) Test these hypotheses using $\alpha = 0.05$. What are your conclusions?

$n = 4$, $\sigma = 3$, $\bar{y} = 1/4\ (145 + 153 + 150 + 147) = 148.75$

$$z_o = \frac{\bar{y} - \mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{148.75 - 150}{\frac{3}{\sqrt{4}}} = \frac{-1.25}{\frac{3}{2}} = -0.8333$$

Since $z_{0.05} = 1.645$, do not reject.

(c) Find the *P*-value for the test in part (b).

From the *z*-table: $P \cong 1 - \left[0.7967 + \left(2/3\right)\left(0.7995 - 0.7967\right)\right] = 0.2014$

(d) Construct a 95 percent confidence interval on the mean breaking strength.

The 95% confidence interval is

$$\bar{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$
$$148.75 - (1.96)(3/2) \le \mu \le 148.75 + (1.96)(3/2)$$

$$145.81 \le \mu \le 151.69$$

2.17. The viscosity of a liquid detergent is supposed to average 800 centistokes at 25°C. A random sample of 16 batches of detergent is collected, and the average viscosity is 812. Suppose we know that the standard deviation of viscosity is $\sigma = 25$ centistokes.

(a) State the hypotheses that should be tested.

H_0: $\mu = 800$ $\quad$ H_1: $\mu \neq 800$

(b) Test these hypotheses using $\alpha = 0.05$. What are your conclusions?

$$z_o = \frac{\bar{y} - \mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{812 - 800}{\frac{25}{\sqrt{16}}} = \frac{12}{\frac{25}{4}} = 1.92$$

Since $z_{\alpha/2} = z_{0.025} = 1.96$, do not reject.

(c) What is the *P*-value for the test?

(d) Find a 95 percent confidence interval on the mean.

The 95% confidence interval is

$$\bar{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$
$$812 - (1.96)(25/4) \le \mu \le 812 + (1.96)(25/4)$$
$$812 - 12.25 \le \mu \le 812 + 12.25$$
$$799.75 \le \mu \le 824.25$$

2.18. The diameters of steel shafts produced by a certain manufacturing process should have a mean diameter of 0.255 inches. The diameter is known to have a standard deviation of $\sigma = 0.0001$ inch. A random sample of 10 shafts has an average diameter of 0.2545 inches.

(a) Set up the appropriate hypotheses on the mean μ.

H_0: $\mu = 0.255$ $\quad$ H_1: $\mu \neq 0.255$

(b) Test these hypotheses using $\alpha = 0.05$. What are your conclusions?

$$n = 10, \quad \sigma = 0.0001, \quad \bar{y} = 0.2545$$

$$z_o = \frac{\bar{y} - \mu_o}{\frac{\sigma}{\sqrt{n}}} = \frac{0.2545 - 0.255}{\frac{0.0001}{\sqrt{10}}} = -15.81$$

Since $z_{0.025} = 1.96$, reject H_0.

(c) Find the *P*-value for this test. $P = 2.6547\text{x}10^{-56}$

(d) Construct a 95 percent confidence interval on the mean shaft diameter.

The 95% confidence interval is

$$\bar{y} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{y} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$0.2545 - (1.96)\left(\frac{0.0001}{\sqrt{10}}\right) \le \mu \le 0.2545 + (1.96)\left(\frac{0.0001}{\sqrt{10}}\right)$$

$$0.254438 \le \mu \le 0.254562$$

2.19. A normally distributed random variable has an unknown mean μ and a known variance $\sigma^2 = 9$. Find the sample size required to construct a 95 percent confidence interval on the mean that has total length of 1.0.

Since $y \sim N(\mu,9)$, a 95% two-sided confidence interval on μ is

If the total interval is to have width 1.0, then the half-interval is 0.5. Since $z_{\alpha/2} = z_{0.025} = 1.96$,

$$(1.96)(3/\sqrt{n}) = 0.5$$
$$\sqrt{n} = (1.96)(3/0.5) = 11.76$$
$$n = (11.76)^2 = 138.30 \cong 139$$

2.20. The shelf life of a carbonated beverage is of interest. Ten bottles are randomly selected and tested, and the following results are obtained:

Days	
108	138
124	163
124	159
106	134
115	139

(a) We would like to demonstrate that the mean shelf life exceeds 120 days. Set up appropriate hypotheses for investigating this claim.

H_0: $\mu = 120$ $\quad\quad$ H_1: $\mu > 120$

(b) Test these hypotheses using $\alpha = 0.01$. What are your conclusions?

$\bar{y} = 131$
$S^2 = 3438 / 9 = 382$
$S = \sqrt{382} = 19.54$

$$t_0 = \frac{\bar{y} - \mu_0}{S/\sqrt{n}} = \frac{131 - 120}{19.54/\sqrt{10}} = 1.78$$

since $t_{0.01,9} = 2.821$; do not reject H_0

Minitab Output

T-Test of the Mean

```
Test of mu = 120.00 vs mu > 120.00

Variable     N      Mean    StDev   SE Mean        T          P
Shelf Life  10    131.00    19.54      6.18     1.78      0.054
```

T Confidence Intervals

```
Variable     N      Mean    StDev  SE Mean       99.0 % CI
Shelf Life  10    131.00    19.54     6.18  (  110.91,  151.09)
```

(c) Find the *P*-value for the test in part (b). P=0.054

(d) Construct a 99 percent confidence interval on the mean shelf life.

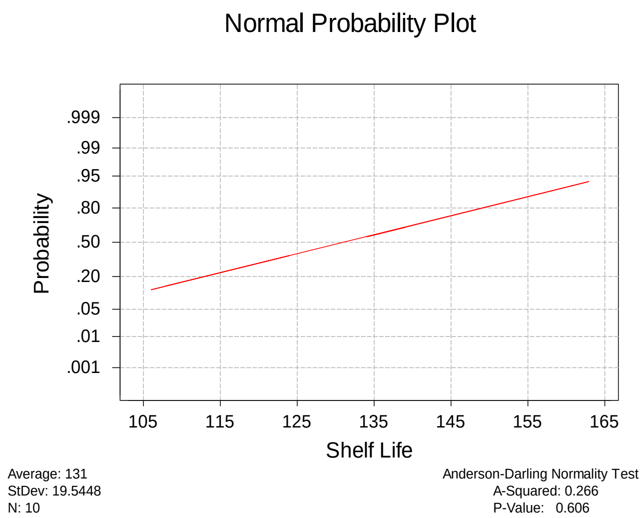
The 99% confidence interval is $\bar{y} - t_{\alpha/2, n-1}\frac{S}{\sqrt{n}} \le \mu \le \bar{y} + t_{\alpha/2, n-1}\frac{S}{\sqrt{n}}$ with $\alpha = 0.01$.

$$131 - (3.250)\left(\frac{19.54}{\sqrt{10}}\right) \le \mu \le 131 + (3.250)\left(\frac{19.54}{\sqrt{10}}\right)$$

$$110.91 \le \mu \le 151.08$$

2.21. Consider the shelf life data in Problem 2.20. Can shelf life be described or modeled adequately by a normal distribution? What effect would violation of this assumption have on the test procedure you used in solving Problem 2.20?

A normal probability plot, obtained from Minitab, is shown. There is no reason to doubt the adequacy of the normality assumption. If shelf life is not normally distributed, then the impact of this on the *t*-test in problem 2.20 is not too serious unless the departure from normality is severe.



2.22. The time to repair an electronic instrument is a normally distributed random variable measured in hours. The repair time for 16 such instruments chosen at random are as follows:

Hours			
159	280	101	212
224	379	179	264
222	362	168	250
149	260	485	170

(a) You wish to know if the mean repair time exceeds 225 hours. Set up appropriate hypotheses for investigating this issue.

$$H_0\text{: } \mu = 225 \qquad H_1\text{: } \mu > 225$$

(b) Test the hypotheses you formulated in part (a). What are your conclusions? Use $\alpha = 0.05$.

$$\bar{y} = 241.50$$
$$S^2 = 146202 / (16 - 1) = 9746.80$$

$$S = \sqrt{9746.8} = 98.73$$

$$t_o = \frac{\bar{y} - \mu_o}{\frac{S}{\sqrt{n}}} = \frac{241.50 - 225}{\frac{98.73}{\sqrt{16}}} = 0.67$$

since $t_{0.05,15} = 1.753$; do not reject H_0

Minitab Output

T-Test of the Mean

```
Test of mu = 225.0 vs mu > 225.0

Variable      N      Mean     StDev   SE Mean        T          P
Hours        16     241.5      98.7      24.7     0.67       0.26
```

```
T Confidence Intervals

Variable     N      Mean     StDev  SE Mean         95.0 % CI
Hours       16     241.5      98.7     24.7   (    188.9,    294.1)
```

(c) Find the P-value for this test. P=0.26

(d) Construct a 95 percent confidence interval on mean repair time.

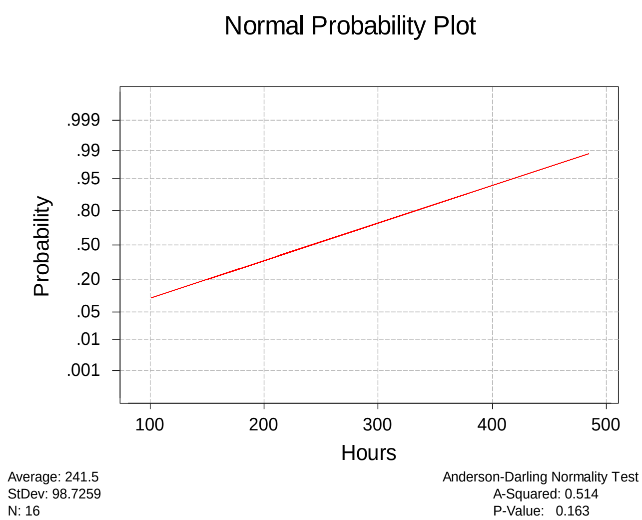
The 95% confidence interval is $\bar{y} - t_{\alpha/2,n-1}\dfrac{S}{\sqrt{n}} \le \mu \le \bar{y} + t_{\alpha/2,n-1}\dfrac{S}{\sqrt{n}}$

$$241.50 - (2.131)\left(\frac{98.73}{\sqrt{16}}\right) \le \mu \le 241.50 + (2.131)\left(\frac{98.73}{\sqrt{16}}\right)$$

$$188.9 \le \mu \le 294.1$$

2.23. Reconsider the repair time data in Problem 2.22. Can repair time, in your opinion, be adequately modeled by a normal distribution?

The normal probability plot below does not reveal any serious problem with the normality assumption.



2.24. Two machines are used for filling plastic bottles with a net volume of 16.0 ounces. The filling processes can be assumed to be normal, with standard deviation of σ_1 = 0.015 and σ_2 = 0.018. The quality engineering department suspects that both machines fill to the same net volume, whether or not this volume is 16.0 ounces. An experiment is performed by taking a random sample from the output of each machine.

Machine 1		Machine 2	
16.03	16.01	16.02	16.03
16.04	15.96	15.97	16.04
16.05	15.98	15.96	16.02
16.05	16.02	16.01	16.01
16.02	15.99	15.99	16.00

(a) State the hypotheses that should be tested in this experiment.

H_0: $\mu_1 = \mu_2$ $\qquad$ H_1: $\mu_1 \neq \mu_2$

(b) Test these hypotheses using α=0.05. What are your conclusions?

$$\bar{y}_1 = 16.015 \qquad \bar{y}_2 = 16.005$$
$$\sigma_1 = 0.015 \qquad \sigma_2 = 0.018$$
$$n_1 = 10 \qquad n_2 = 10$$

$$z_o = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{16.015 - 16.018}{\sqrt{\frac{0.015^2}{10} + \frac{0.018^2}{10}}} = 1.35$$

$z_{0.025}$ = 1.96; do not reject

(c) What is the *P*-value for the test? P = 0.1770

(d) Find a 95 percent confidence interval on the difference in the mean fill volume for the two machines.

The 95% confidence interval is

$$\bar{y}_1 - \bar{y}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \bar{y}_1 - \bar{y}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$
$$(16.015 - 16.005) - (1.96)\sqrt{\frac{0.015^2}{10} + \frac{0.018^2}{10}} \le \mu_1 - \mu_2 \le (16.015 - 16.005) + (1.96)\sqrt{\frac{0.015^2}{10} + \frac{0.018^2}{10}}$$
$$-0.0045 \le \mu_1 - \mu_2 \le 0.0245$$

2.25. Two types of plastic are suitable for use by an electronic calculator manufacturer. The breaking strength of this plastic is important. It is known that $\sigma_1 = \sigma_2$ = 1.0 psi. From random samples of n_1 = 10 and n_2 = 12 we obtain $\bar{y}_1$ = 162.5 and $\bar{y}_2$ = 155.0. The company will not adopt plastic 1 unless its breaking strength exceeds that of plastic 2 by at least 10 psi. Based on the sample information, should they use plastic 1? In answering this questions, set up and test appropriate hypotheses using α = 0.01. Construct a 99 percent confidence interval on the true mean difference in breaking strength.

H_0: $\mu_1 - \mu_2 = 10$ $\qquad$ H_1: $\mu_1 - \mu_2 > 10$

$$\bar{y}_1 = 162.5 \qquad \bar{y}_2 = 155.0$$
$$\sigma_1 = 1 \qquad \sigma_2 = 1$$
$$n_1 = 10 \qquad n_2 = 10$$

$$z_o = \frac{\bar{y}_1 - \bar{y}_2 - 10}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{162.5 - 155.0 - 10}{\sqrt{\frac{1^2}{10} + \frac{1^2}{12}}} = -5.84$$

$z_{0.01}$ = 2.325; do not reject

The 99 percent confidence interval is

$$\bar{y}_1 - \bar{y}_2 - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \le \mu_1 - \mu_2 \le \bar{y}_1 - \bar{y}_2 + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$

$$(162.5-155.0)-(2.575)\sqrt{\frac{1^2}{10}+\frac{1^2}{12}} \le \mu_1 - \mu_2 \le (162.5-155.0)+(2.575)\sqrt{\frac{1^2}{10}+\frac{1^2}{12}}$$

$$6.40 \le \mu_1 - \mu_2 \le 8.60$$

2.26. The following are the burning times (in minutes) of chemical flares of two different formulations. The design engineers are interested in both the means and variance of the burning times.

Type 1		Type 2	
65	82	64	56
81	67	71	69
57	59	83	74
66	75	59	82
82	70	65	79

(a) Test the hypotheses that the two variances are equal. Use $\alpha = 0.05$.

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$

Do not reject.

(b) Using the results of (a), test the hypotheses that the mean burning times are equal. Use $\alpha = 0.05$. What is the *P*-value for this test?

Do not reject.

From the computer output, t=0.05; do not reject. Also from the computer output P=0.96

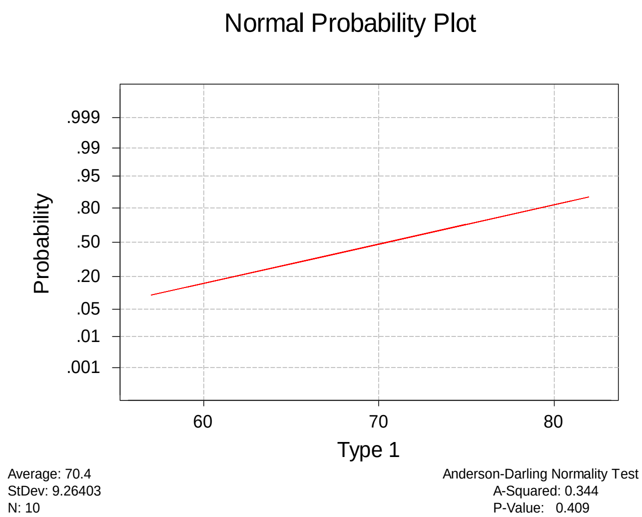
Minitab Output

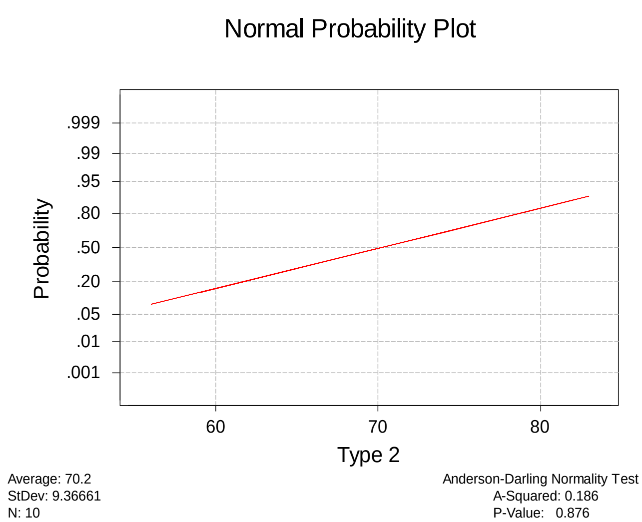
Two Sample T-Test and Confidence Interval

```
Two sample T for Type 1 vs Type 2

         N      Mean     StDev   SE Mean
Type 1  10     70.40      9.26       2.9
Type 2  10     70.20      9.37       3.0

95% CI for mu Type 1 - mu Type 2: ( -8.6,  9.0)
T-Test mu Type 1 = mu Type 2 (vs not =): T = 0.05  P = 0.96  DF = 18
Both use Pooled StDev = 9.32
```

(c) Discuss the role of the normality assumption in this problem. Check the assumption of normality for both types of flares.

The assumption of normality is required in the theoretical development of the *t*-test. However, moderate departure from normality has little impact on the performance of the *t*-test. The normality assumption is more important for the test on the equality of the two variances. An indication of nonnormality would be of concern here. The normal probability plots shown below indicate that burning time for both formulations follow the normal distribution.





2.27. An article in *Solid State Technology*, "Orthogonal Design of Process Optimization and Its Application to Plasma Etching" by G.Z. Yin and D.W. Jillie (May, 1987) describes an experiment to determine the effect of C_2F_6 flow rate on the uniformity of the etch on a silicon wafer used in integrated circuit manufacturing. Data for two flow rates are as follows:

C_2F_6	Uniformity Observation					
(SCCM)	1	2	3	4	5	6
125	2.7	4.6	2.6	3.0	3.2	3.8
200	4.6	3.4	2.9	3.5	4.1	5.1

(a) Does the C_2F_6 flow rate affect average etch uniformity? Use $\alpha = 0.05$.

No, C_2F_6 flow rate does not affect average etch uniformity.

Minitab Output

```
Two Sample T-Test and Confidence Interval

Two sample T for Uniformity

Flow Rat    N      Mean     StDev   SE Mean
125         6     3.317     0.760      0.31
200         6     3.933     0.821      0.34

95% CI for mu (125) - mu (200): ( -1.63,  0.40)
T-Test mu (125) = mu (200) (vs not =): T = -1.35  P = 0.21  DF = 10
Both use Pooled StDev = 0.791
```

(b) What is the *P*-value for the test in part (a)? From the *Minitab* output, *P*=0.21

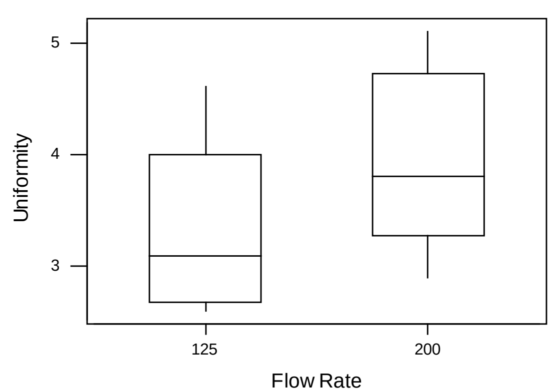
(c) Does the C_2F_6 flow rate affect the wafer-to-wafer variability in etch uniformity? Use $\alpha = 0.05$.

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$
$$F_{0.025,5,5} = 7.15$$
$$F_{0.975,5,5} = 0.14$$
$$F_0 = \frac{0.5776}{0.6724} = 0.86$$

Do not reject; C_2F_6 flow rate does not affect wafer-to-wafer variability.

(d) Draw box plots to assist in the interpretation of the data from this experiment.

The box plots shown below indicate that there is little difference in uniformity at the two gas flow rates. Any observed difference is not statistically significant. See the *t*-test in part (a).



2.28. A new filtering device is installed in a chemical unit. Before its installation, a random sample yielded the following information about the percentage of impurity: $\bar{y}_1 = 12.5$, $S_1^2 = 101.17$, and $n_1 = 8$. After installation, a random sample yielded $\bar{y}_2 = 10.2$, $S_2^2 = 94.73$, $n_2 = 9$.

(a) Can you conclude that the two variances are equal? Use $\alpha = 0.05$.

$$H_0 : \sigma_1^2 = \sigma_2^2$$
$$H_1 : \sigma_1^2 \neq \sigma_2^2$$
$$F_{0.025,7,8} = 4.53$$
$$F_0 = \frac{S_1^2}{S_2^2} = \frac{101.17}{94.73} = 1.07$$

Do not reject. Assume that the variances are equal.

(b) Has the filtering device reduced the percentage of impurity significantly? Use $\alpha = 0.05$.

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 > \mu_2$$
$$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2} = \frac{(8-1)(101.17)+(9-1)(94.73)}{8+9-2} = 97.74$$
$$S_p = 9.89$$
$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{12.5-10.2}{9.89\sqrt{\frac{1}{8}+\frac{1}{9}}} = 0.479$$
$$t_{0.05,15} = 1.753$$

Do not reject. There is no evidence to indicate that the new filtering device has affected the mean.

2.29. Photoresist is a light-sensitive material applied to semiconductor wafers so that the circuit pattern can be imaged on to the wafer. After application, the coated wafers are baked to remove the solvent in the photoresist mixture and to harden the resist. Here are measurements of photoresist thickness (in kÅ) for eight wafers baked at two different temperatures. Assume that all of the runs were made in random order.

95 ºC	100 ºC
11.176	5.623
7.089	6.748
8.097	7.461
11.739	7.015
11.291	8.133
10.759	7.418
6.467	3.772
8.315	8.963

(a) Is there evidence to support the claim that the higher baking temperature results in wafers with a lower mean photoresist thickness? Use $\alpha = 0.05$.

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 > \mu_2$$
$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{(8-1)(4.41) + (8-1)(2.54)}{8+8-2} = 3.48$$
$$S_p = 1.86$$
$$t_0 = \frac{\overline{y}_1 - \overline{y}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{9.37 - 6.89}{1.86\sqrt{\frac{1}{8} + \frac{1}{8}}} = 2.65$$
$$t_{0.05,14} = 1.761$$

Since $t_{0.05,14} = 1.761$, reject H_0. There appears to be a lower mean thickness at the higher temperature. This is also seen in the computer output.

Minitab Output

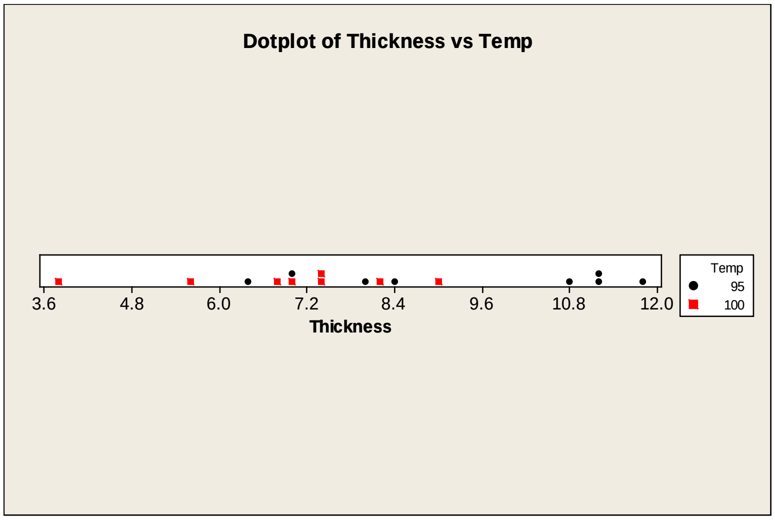
```
Two-Sample T-Test and CI: Thickness, Temp

Two-sample T for Thick@95 vs Thick@100

          N      Mean     StDev   SE Mean
Thick@95  8      9.37      2.10      0.74
Thick@10  8      6.89      1.60      0.56

Difference = mu Thick@95 - mu Thick@100
Estimate for difference:  2.475
95% lower bound for difference: 0.833
T-Test of difference = 0 (vs >): T-Value = 2.65  P-Value = 0.009  DF = 14
Both use Pooled StDev = 1.86
```

(b) What is the *P*-value for the test conducted in part (a)? $P = 0.009$

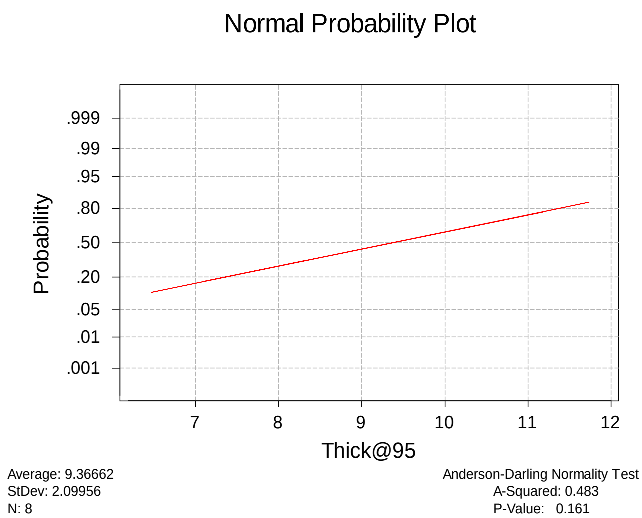
(c) Find a 95% confidence interval on the difference in means. Provide a practical interpretation of this interval.

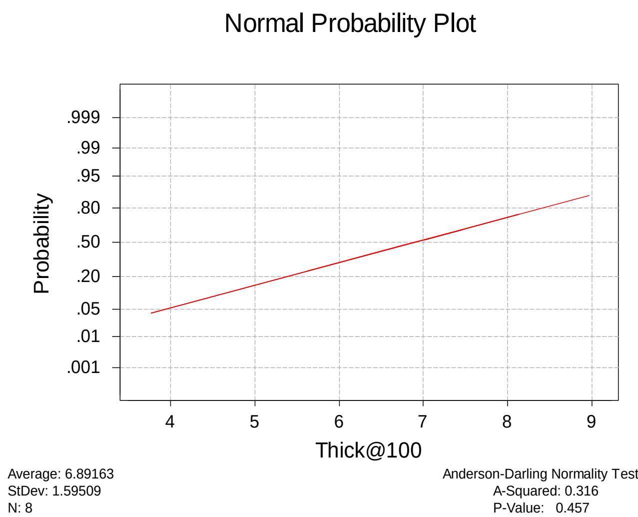
From the computer output the 95% lower confidence bound is $0.833 \le \mu_1 - \mu_2$. This lower confidence bound is greater than 0; therefore, there is a difference in the two temperatures on the thickness of the photoresist.

(d) Draw dot diagrams to assist in interpreting the results from this experiment.



(e) Check the assumption of normality of the photoresist thickness.





There are no significant deviations from the normality assumptions.

(f) Find the power of this test for detecting an actual difference in means of 2.5 kÅ.

Minitab Output
```
Power and Sample Size

2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Sigma = 1.86

            Sample
Difference    Size   Power
       2.5       8  0.7056
```

(g) What sample size would be necessary to detect an actual difference in means of 1.5 kÅ with a power of at least 0.9?.

Minitab Output
```
Power and Sample Size

2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Sigma = 1.86

            Sample  Target  Actual
Difference    Size   Power   Power
       1.5      34  0.9000  0.9060
```

This result makes intuitive sense. More samples are needed to detect a smaller difference.

2.30. Front housings for cell phones are manufactured in an injection molding process. The time the part is allowed to cool in the mold before removal is thought to influence the occurrence of a particularly troublesome cosmetic defect, flow lines, in the finished housing. After manufacturing, the housings are inspected visually and assigned a score between 1 and 10 based on their appearance, with 10 corresponding to a perfect part and 1 corresponding to a completely defective part. An experiment was conducted using two cool-down times, 10 seconds and 20 seconds, and 20 housings were evaluated at each level of cool-down time. All 40 observations in this experiment were run in random order. The data are shown below.

10 Seconds		20 Seconds	
1	3	7	6
2	6	8	9
1	5	5	5
3	3	9	7
5	2	5	4
1	1	8	6
5	6	6	8
2	8	4	5
3	2	6	8
5	3	7	7

(a) Is there evidence to support the claim that the longer cool-down time results in fewer appearance defects? Use $\alpha = 0.05$.

From the analysis shown below, there is evidence that the longer cool-down time results in fewer appearance defects.

Minitab Output

Two-Sample T-Test and CI: 10 seconds, 20 seconds

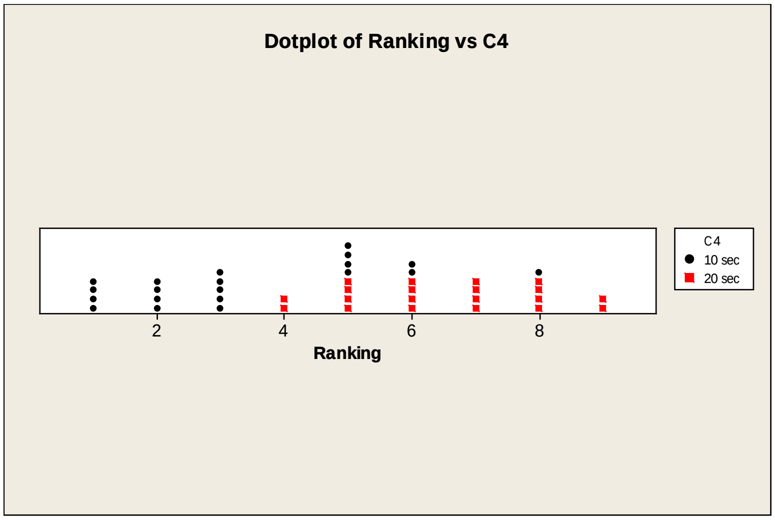
```
Two-sample T for 10 seconds vs 20 seconds

           N      Mean     StDev   SE Mean
10 secon  20      3.35      2.01      0.45
20 secon  20      6.50      1.54      0.34

Difference = mu 10 seconds - mu 20 seconds
Estimate for difference:  -3.150
95% upper bound for difference: -2.196
T-Test of difference = 0 (vs <): T-Value = -5.57  P-Value = 0.000  DF = 38
Both use Pooled StDev = 1.79
```

(b) What is the P-value for the test conducted in part (a)? From the *Minitab* output, $P = 0.000$

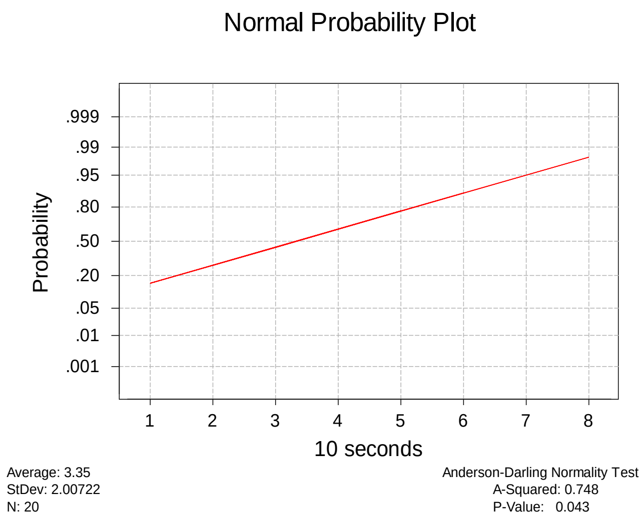
(c) Find a 95% confidence interval on the difference in means. Provide a practical interpretation of this interval.

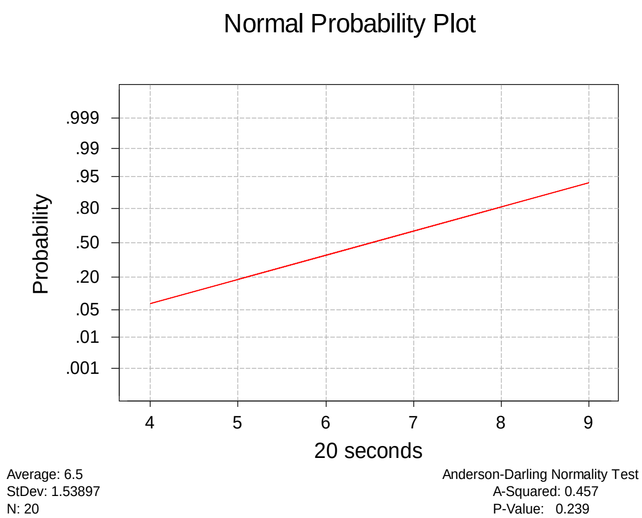
From the *Minitab* output, $\mu_1 - \mu_2 \leq -2.196$. This lower confidence bound is less than 0. The two samples are different. The 20 second cooling time gives a cosmetically better housing.

(d) Draw dot diagrams to assist in interpreting the results from this experiment.



(e) Check the assumption of normality for the data from this experiment.





There are no significant departures from normality.

2.31. Twenty observations on etch uniformity on silicon wafers are taken during a qualification experiment for a plasma etcher. The data are as follows:

Etch Uniformity				
5.34	6.65	4.76	5.98	7.25
6.00	7.55	5.54	5.62	6.21
5.97	7.35	5.44	4.39	4.98
5.25	6.35	4.61	6.00	5.32

(a) Construct a 95 percent confidence interval estimate of σ^2.

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{(1-\alpha/2),n-1}}$$

$$\frac{(20-1)(0.88907)^2}{32.852} \le \sigma^2 \le \frac{(20-1)(0.88907)^2}{8.907}$$

$$0.457 \le \sigma^2 \le 1.686$$

(b) Test the hypothesis that $\sigma^2 = 1.0$. Use $\alpha = 0.05$. What are your conclusions?

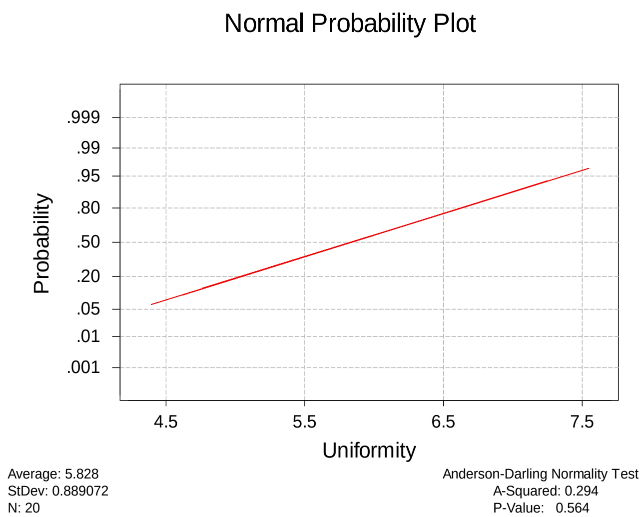
Do not reject. There is no evidence to indicate that $\sigma^2 \ne 1$

(c) Discuss the normality assumption and its role in this problem.

The normality assumption is much more important when analyzing variances then when analyzing means. A moderate departure from normality could cause problems with both statistical tests and confidence intervals. Specifically, it will cause the reported significance levels to be incorrect.

(d) Check normality by constructing a normal probability plot. What are your conclusions?

The normal probability plot indicates that there is not a serious problem with the normality assumption.



2.32. The diameter of a ball bearing was measured by 12 inspectors, each using two different kinds of calipers. The results were:

Inspector	Caliper 1	Caliper 2	Difference	Difference^2
1	0.265	0.264	0.001	0.000001
2	0.265	0.265	0.000	0
3	0.266	0.264	0.002	0.000004
4	0.267	0.266	0.001	0.000001
5	0.267	0.267	0.000	0
6	0.265	0.268	-0.003	0.000009
7	0.267	0.264	0.003	0.000009
8	0.267	0.265	0.002	0.000004
9	0.265	0.265	0.000	0
10	0.268	0.267	0.001	0.000001
11	0.268	0.268	0.000	0
12	0.265	0.269	-0.004	0.000016
			$\sum = 0.003$	$\sum = 0.000045$

(a) Is there a significant difference between the means of the population of measurements represented by the two samples? Use $\alpha = 0.05$.

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2 \end{aligned} \quad \text{or equivalently} \quad \begin{aligned} H_0 &: \mu_d = 0 \\ H_1 &: \mu_d \neq 0 \end{aligned}$$

Minitab Output

Paired T-Test and Confidence Interval

```
Paired T for Caliper 1 - Caliper 2

                  N      Mean     StDev   SE Mean
Caliper          12  0.266250  0.001215  0.000351
Caliper          12  0.266000  0.001758  0.000508
Difference       12  0.000250  0.002006  0.000579

95% CI for mean difference: (-0.001024, 0.001524)
T-Test of mean difference = 0 (vs not = 0): T-Value = 0.43  P-Value = 0.674
```

(b) Find the *P*-value for the test in part (a). *P*=0.674

(c) Construct a 95 percent confidence interval on the difference in the mean diameter measurements for the two types of calipers.

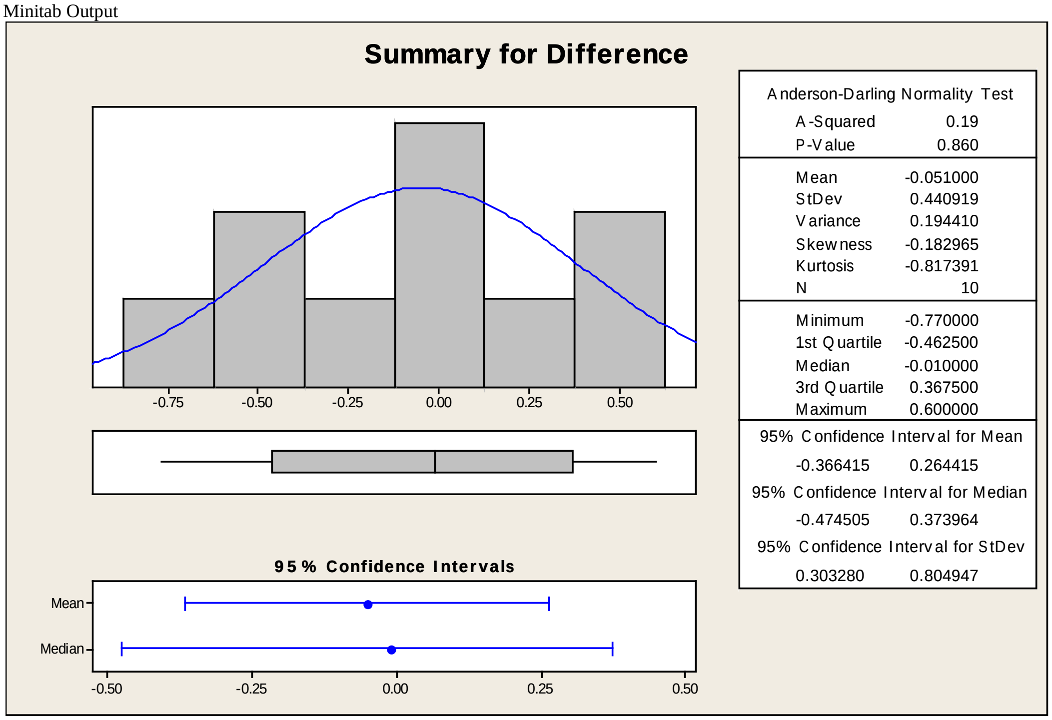
$$\bar{d} - t_{\alpha/2,n-1}\frac{S_d}{\sqrt{n}} \le \mu_D \left(= \mu_1 - \mu_2\right) \le \bar{d} + t_{\alpha/2,n-1}\frac{S_d}{\sqrt{n}}$$
$$0.00025 - 2.201\frac{0.002}{\sqrt{12}} \le \mu_d \le 0.00025 + 2.201\frac{0.002}{\sqrt{12}}$$
$$-0.00102 \le \mu_d \le 0.00152$$

2.33. An article in the journal of *Neurology* (1998, Vol. 50, pp.1246-1252) observed that the monozygotic twins share numerous physical, psychological and pathological traits. The investigators measured an intelligence score of 10 pairs of twins. The data are obtained as follows:

Pair	Birth Order: 1	Birth Order: 2
1	6.08	5.73
2	6.22	5.80
3	7.99	8.42
4	7.44	6.84
5	6.48	6.43
6	7.99	8.76
7	6.32	6.32
8	7.60	7.62
9	6.03	6.59
10	7.52	7.67

(a) Is the assumption that the difference in score is normally distributed reasonable?

Minitab Output



By plotting the differences, the output shows that the Anderson-Darling Normality Test shows a P-Value of 0.860. The data is assumed to be normal.

(b) Find a 95% confidence interval on the difference in the mean score. Is there any evidence that mean score depends on birth order?

The 95% confidence interval on the difference in mean score is (-0.366415, 0.264415) contains the value of zero. There is no difference in birth order.

(c) Test an appropriate set of hypothesis indicating that the mean score does not depend on birth order.

$$H_0: \mu_1 = \mu_2 \quad \text{or equivalently} \quad H_0: \mu_d = 0$$
$$H_1: \mu_1 \neq \mu_2 \qquad\qquad\qquad\quad H_1: \mu_d \neq 0$$

Minitab Output

```
Paired T for Birth Order: 1 - Birth Order: 2

                 N    Mean  StDev  SE Mean
Birth Order: 1  10   6.967  0.810    0.256
Birth Order: 2  10   7.018  1.053    0.333
Difference      10  -0.051  0.441    0.139


95% CI for mean difference: (-0.366, 0.264)
T-Test of mean difference = 0 (vs not = 0): T-Value = -0.37  P-Value = 0.723
```

Do not reject. The *P*-value is 0.723.

2.34. An article in the *Journal of Strain Analysis* (vol.18, no. 2, 1983) compares several procedures for predicting the shear strength for steel plate girders. Data for nine girders in the form of the ratio of predicted to observed load for two of these procedures, the Karlsruhe and Lehigh methods, are as follows:

Girder	Karlsruhe Method	Lehigh Method	Difference	Difference^2
S1/1	1.186	1.061	0.125	0.015625
S2/1	1.151	0.992	0.159	0.025281
S3/1	1.322	1.063	0.259	0.067081
S4/1	1.339	1.062	0.277	0.076729
S5/1	1.200	1.065	0.135	0.018225
S2/1	1.402	1.178	0.224	0.050176
S2/2	1.365	1.037	0.328	0.107584
S2/3	1.537	1.086	0.451	0.203401
S2/4	1.559	1.052	0.507	0.257049
		Sum =	2.465	0.821151
		Average =	0.274	

(a) Is there any evidence to support a claim that there is a difference in mean performance between the two methods? Use $\alpha = 0.05$.

$$H_0: \mu_1 = \mu_2 \quad \text{or equivalently} \quad H_0: \mu_d = 0$$
$$H_1: \mu_1 \neq \mu_2 \qquad\qquad\qquad\quad H_1: \mu_d \neq 0$$

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n} d_i = \frac{1}{9}(2.465) = 0.274$$

$$s_d = \left[\frac{\sum_{i=1}^{n} d_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} d_i\right)^2}{n-1}\right]^{\frac{1}{2}} = \left[\frac{0.821151 - \frac{1}{9}(2.465)^2}{9-1}\right]^{\frac{1}{2}} = 0.135$$

$$t_0 = \frac{\bar{d}}{\frac{S_d}{\sqrt{n}}} = \frac{0.274}{\frac{0.135}{\sqrt{9}}} = 6.08$$

$t_{\alpha/2, n-1} = t_{0.025,8} = 2.306$, reject the null hypothesis.

Minitab Output

Paired T-Test and Confidence Interval

Paired T for Karlsruhe - Lehigh

	N	Mean	StDev	SE Mean
Karlsruh	9	1.3401	0.1460	0.0487
Lehigh	9	1.0662	0.0494	0.0165
Difference	9	0.2739	0.1351	0.0450

95% CI for mean difference: (0.1700, 0.3777)
T-Test of mean difference = 0 (vs not = 0): T-Value = 6.08 P-Value = 0.000

(b) What is the *P*-value for the test in part (a)?

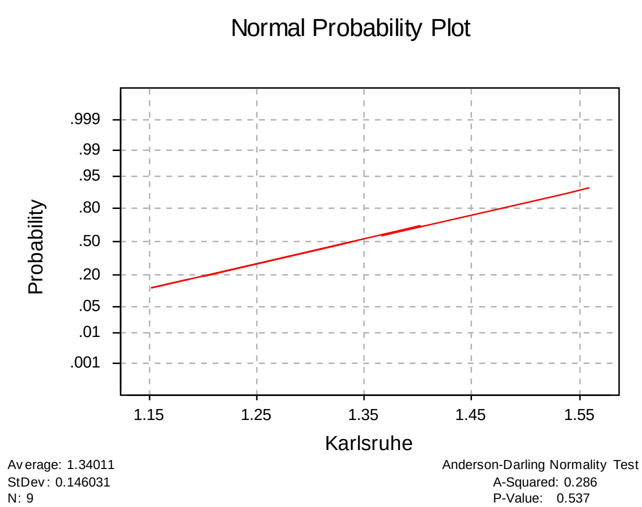
P=0.0002

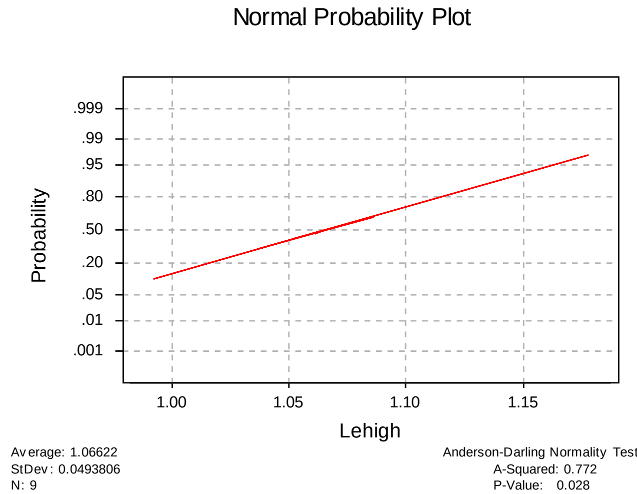
(c) Construct a 95 percent confidence interval for the difference in mean predicted to observed load.

$$\bar{d} - t_{\alpha/2, n-1}\frac{S_d}{\sqrt{n}} \le \mu_d \le \bar{d} + t_{\alpha/2, n-1}\frac{S_d}{\sqrt{n}}$$
$$0.274 - 2.306\frac{0.135}{\sqrt{9}} \le \mu_d \le 0.274 + 2.306\frac{0.135}{\sqrt{9}}$$
$$0.17023 \le \mu_d \le 0.37777$$

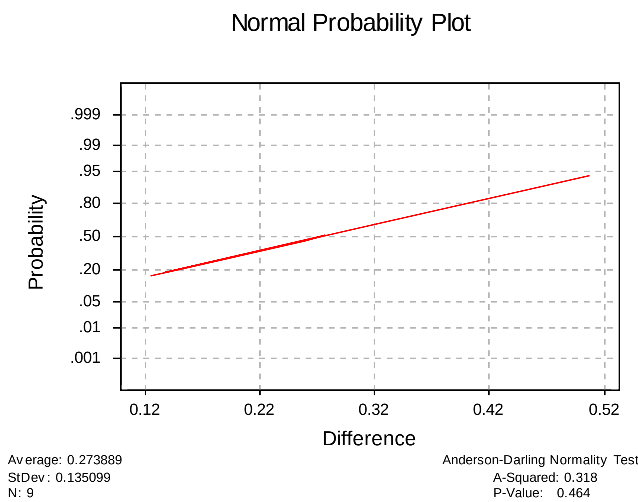
(d) Investigate the normality assumption for both samples.

The normal probability plots of the observations for each method follow. There are no serious concerns with the normality assumption, but there is an indication of a possible outlier (1.178) in the Lehigh method data.





(e) Investigate the normality assumption for the difference in ratios for the two methods.



There is no issue with normality in the difference of ratios of the two methods.

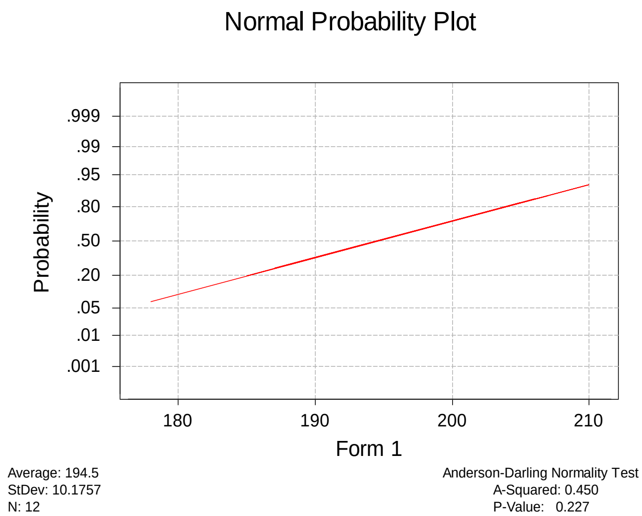
(f) Discuss the role of the normality assumption in the paired *t*-test.

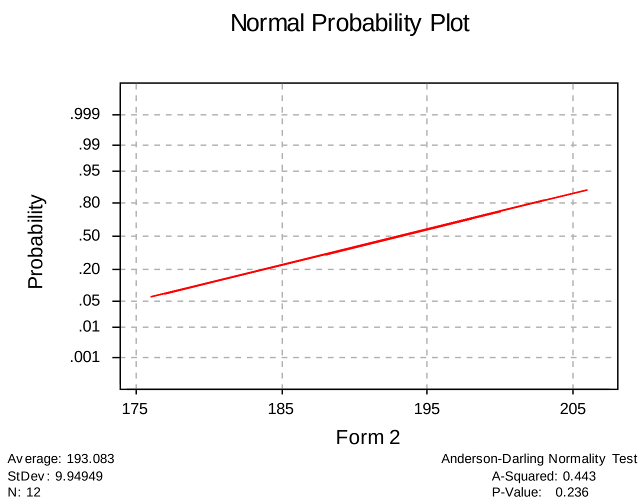
As in any *t*-test, the assumption of normality is of only moderate importance. In the paired *t*-test, the assumption of normality applies to the distribution of the differences. That is, the individual sample measurements do not have to be normally distributed, only their difference.

2.35. The deflection temperature under load for two different formulations of ABS plastic pipe is being studied. Two samples of 12 observations each are prepared using each formulation, and the deflection temperatures (in °F) are reported below:

Formulation 1			Formulation 2		
206	193	192	177	176	198
188	207	210	197	185	188
205	185	194	206	200	189
187	189	178	201	197	203

(a) Construct normal probability plots for both samples. Do these plots support assumptions of normality and equal variance for both samples?





(b) Do the data support the claim that the mean deflection temperature under load for formulation 1 exceeds that of formulation 2? Use $\alpha = 0.05$.

No, formulation 1 does not exceed formulation 2 per the *Minitab* output below.

Minitab Output

Two Sample T-Test and Confidence Interval

```
         N      Mean     StDev   SE Mean
Form 1  12     194.5      10.2       2.9
Form 2  12    193.08      9.95       2.9

Difference = mu Form 1 - mu Form 2
Estimate for difference:  1.42
95% lower bound for difference: -5.64
T-Test of difference = 0 (vs >): T-Value = 0.34  P-Value = 0.367  DF = 22
Both use Pooled StDev = 10.1
```

(c) What is the *P*-value for the test in part (a)?

$P = 0.367$

2.36. Refer to the data in problem 2.35. Do the data support a claim that the mean deflection temperature under load for formulation 1 exceeds that of formulation 2 by at least 3 °F?

No, formulation 1 does not exceed formulation 2 by at least 3 °F.

Minitab Output

Two-Sample T-Test and CI: Form1, Form2

```
Two-sample T for Form 1 vs Form 2

         N      Mean     StDev   SE Mean
Form 1  12     194.5      10.2       2.9
Form 2  12    193.08      9.95       2.9

Difference = mu Form 1 - mu Form 2
Estimate for difference:  1.42
95% lower bound for difference: -5.64
T-Test of difference = 3 (vs >): T-Value = -0.39  P-Value = 0.648  DF = 22
Both use Pooled StDev = 10.1
```

2.37. In semiconductor manufacturing, wet chemical etching is often used to remove silicon from the backs of wafers prior to metalization. The etch rate is an important characteristic of this process. Two different etching solutions are being evaluated. Eight randomly selected wafers have been etched in each solution and the observed etch rates (in mils/min) are shown below:

Solution 1		Solution 2	
9.9	10.6	10.2	10.6
9.4	10.3	10.0	10.2
10.0	9.3	10.7	10.4
10.3	9.8	10.5	10.3

(a) Do the data indicate that the claim that both solutions have the same mean etch rate is valid? Use $\alpha = 0.05$ and assume equal variances.

No, the solutions do not have the same mean etch rate. See the Minitab output below.

Minitab Output

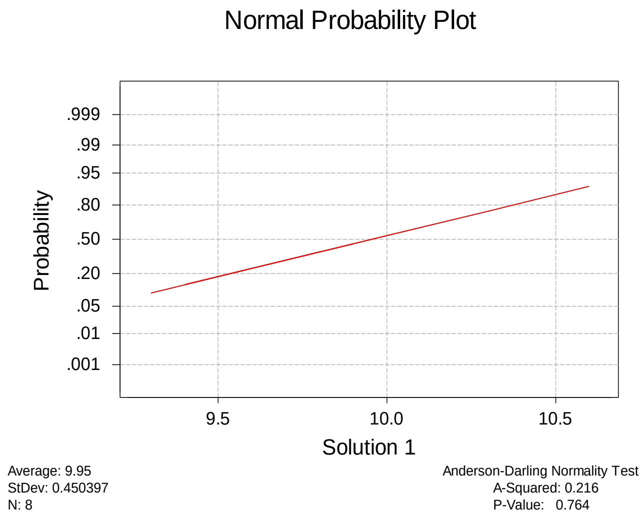
Two Sample T-Test and Confidence Interval

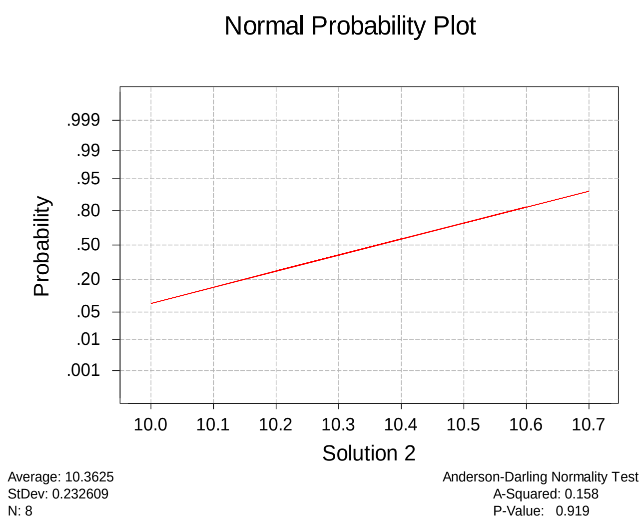
```
Two-sample T for Solution 1 vs Solution 2

          N      Mean     StDev   SE Mean
Solution  8     9.950     0.450      0.16
Solution  8    10.363     0.233     0.082

Difference = mu Solution 1 - mu Solution 2
Estimate for difference:  -0.413
95% CI for difference: (-0.797, -0.028)
T-Test of difference = 0 (vs not =): T-Value = -2.30  P-Value = 0.037  DF = 14
Both use Pooled StDev = 0.358
```

(b) Find a 95% confidence interval on the difference in mean etch rate.

From the Minitab output, -0.797 to –0.028.

(c) Use normal probability plots to investigate the adequacy of the assumptions of normality and equal variances.





Both the normality and equality of variance assumptions are valid.

2.38. Two popular pain medications are being compared on the basis of the speed of absorption by the body. Specifically, tablet 1 is claimed to be absorbed twice as fast as tablet 2. Assume that σ_1^2 and σ_2^2 are known. Develop a test statistic for

$$H_0\text{: } 2\mu_1 = \mu_2$$
$$H_1\text{: } 2\mu_1 \neq \mu_2$$

$2\bar{y}_1 - \bar{y}_2 \sim N\left(2\mu_1 - \mu_2, \dfrac{4\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}\right)$, assuming that the data is normally distributed.

The test statistic is: $z_o = \dfrac{2\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{4\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$, reject if $|z_o| > z_{\alpha/2}$

2.39. Continuation of Problem 2.38. An article in *Nature* (1972, pp.225-226) reported on the levels of monoamine oxidase in blood platelets for a sample of 43 schizophrenic patients resulting in $\bar{y}_1$ = 2.69 and s_1 = 2.30 while for a sample of 45 normal patients the results were $\bar{y}_2$ = 6.35 and s_2 = 4.03. The units are nm/mg protein/h. Use the results of the previous problem to test the claim that the mean monoamine oxidase level for normal patients is at least twice the mean level for schizophrenic patients. Assume that the sample sizes are large enough to use the sample standard deviations as the true parameter values.

$$z_o = \frac{2\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{4\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{2(2.69) - 6.35}{\sqrt{\dfrac{4(2.30)^2}{43} + \dfrac{4.03^2}{45}}} = \frac{-0.97}{.92357} = -1.05$$

z_0 = -1.05; using α=0.05, $z_{\alpha/2} = 1.96$, do not reject.

2.40. Suppose we are testing

H_0: $\mu_1 = \mu_2$
H_1: $\mu_1 \neq \mu_2$

where σ_1^2 and σ_2^2 are known. Our sampling resources are constrained such that $n_1 + n_2 = N$. How should we allocate the n_1, n_2 to the two samples that lead to the most powerful test?

The most powerful test is attained by the n_1 and n_2 that maximize z_o for given $\bar{y}_1 - \bar{y}_2$.
Thus, we chose n_1 and n_2 to $\max z_o = \dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$, subject to $n_1 + n_2 = N$.

This is equivalent to min $L = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2} = \dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{N - n_1}$, subject to $n_1 + n_2 = N$.

Now $\dfrac{dL}{dn_1} = \dfrac{-\sigma_1^2}{n_1^2} + \dfrac{\sigma_2^2}{(N - n_1)^2} = 0$, implies that $n_1 / n_2 = \sigma_1 / \sigma_2$.

Thus n_1 and n_2 are assigned proportionally to the ratio of the standard deviations. This has intuitive appeal, as it allocates more observations to the population with the greatest variability.

2.41 Continuation of Problem 2.40. Suppose that we want to construct a 95% two-sided confidence interval on the difference in two means where the two sample standard deviations are known to be σ_1 = 4 and σ_2 = 8. The total sample size is restricted to N = 30. What is the length of the 95% CI if the sample sizes used by the experimenter are $n_1 = n_2 = 15$? How much shorter would the 95% CI have been if the experiment had used the optimal sample size calculation?

The 95% confidence interval for $n_1 = n_2 = 15$ is

$$(\bar{y}_1-\bar{y}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\le\mu_1-\mu_2\le(\bar{y}_1-\bar{y}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$
$$(\bar{y}_1-\bar{y}_2)-z_{\alpha/2}\sqrt{\frac{4^2}{15}+\frac{8^2}{15}}\le\mu_1-\mu_2\le(\bar{y}_1-\bar{y}_2)+z_{\alpha/2}\sqrt{\frac{4^2}{15}+\frac{8^2}{15}}$$
$$(\bar{y}_1-\bar{y}_2)-z_{\alpha/2}(2.31)\le\mu_1-\mu_2\le(\bar{y}_1-\bar{y}_2)+z_{\alpha/2}(2.31)$$

The 95% confidence interval for the proportions is,

$$n_1 = 30 - n_2$$

$$\frac{n_1}{n_2}=\frac{\sigma_1}{\sigma_2}=\frac{30-n_2}{n_2}=\frac{4}{8}$$

Therefore n_2 = 20 and n_1 = 10

$$(\bar{y}_1-\bar{y}_2)-z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\le\mu_1-\mu_2\le(\bar{y}_1-\bar{y}_2)+z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$$
$$(\bar{y}_1-\bar{y}_2)-z_{\alpha/2}\sqrt{\frac{4^2}{10}+\frac{8^2}{20}}\le\mu_1-\mu_2\le(\bar{y}_1-\bar{y}_2)+z_{\alpha/2}\sqrt{\frac{4^2}{10}+\frac{8^2}{20}}$$
$$(\bar{y}_1-\bar{y}_2)-z_{\alpha/2}(2.19)\le\mu_1-\mu_2\le(\bar{y}_1-\bar{y}_2)+z_{\alpha/2}(2.19)$$

The confidence interval decreases from a multiple of 2.31 to a multiple of 2.19.

2.42. Develop Equation 2.46 for a 100(1 - α) percent confidence interval for the variance of a normal distribution.

$\frac{SS}{\sigma^2}\sim\chi^2_{n-1}$. Thus, $P\left\{\chi^2_{1-\frac{\alpha}{2},n-1}\le\frac{SS}{\sigma^2}\le\chi^2_{\frac{\alpha}{2},n-1}\right\}=1-\alpha$. Therefore,

$$P\left\{\frac{SS}{\chi^2_{\frac{\alpha}{2},n-1}}\le\sigma^2\le\frac{SS}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right\}=1-\alpha,$$

so $\left[\frac{SS}{\chi^2_{\frac{\alpha}{2},n-1}},\frac{SS}{\chi^2_{1-\frac{\alpha}{2},n-1}}\right]$ is the 100(1 - α)% confidence interval on σ^2.

2.43. Develop Equation 2.50 for a 100(1 - α) percent confidence interval for the ratio σ_1^2 / σ_2^2, where σ_1^2 and σ_2^2 are the variances of two normal distributions.

$$\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}\sim F_{n_2-1,n_1-1}$$

$$P\left\{F_{1-\alpha/2,n_2-1,n_1-1}\le\frac{S_2^2/\sigma_2^2}{S_1^2/\sigma_1^2}\le F_{\alpha/2,n_2-1,n_1-1}\right\}=1-\alpha \quad \text{or}$$

$$P\left\{\frac{S_1^2}{S_2^2}F_{1-\alpha/2,n_2-1,n_1-1}\le\frac{\sigma_1^2}{\sigma_2^2}\le\frac{S_1^2}{S_2^2}F_{\alpha/2,n_2-1,n_1-1}\right\}=1-\alpha$$

2.44. Develop an equation for finding a 100(1 - α) percent confidence interval on the difference in the means of two normal distributions where $\sigma_1^2 \neq \sigma_2^2$. Apply your equation to the portland cement experiment data, and find a 95% confidence interval.

$$\frac{(\bar{y}_1-\bar{y}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}} \sim t_{\alpha/2,\upsilon}$$

$$t_{\alpha/2,\upsilon}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \le (\bar{y}_1-\bar{y}_2)-(\mu_1-\mu_2) \le t_{\alpha/2,\upsilon}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

$$(\bar{y}_1-\bar{y}_2)-t_{\alpha/2,\upsilon}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}} \le (\mu_1-\mu_2) \le (\bar{y}_1-\bar{y}_2)+t_{\alpha/2,\upsilon}\sqrt{\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}}$$

$$\text{where}\quad \upsilon = \frac{\left(\frac{S_1^2}{n_1}+\frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1-1}+\frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2-1}}$$

Using the data from Table 2.1

$$(16.764-17.343)-2.110\sqrt{\frac{0.100138}{10}+\frac{0.0614622}{10}} \le (\mu_1-\mu_2) \le$$
$$(16.764-17.343)+2.110\sqrt{\frac{0.100138}{10}+\frac{0.0614622}{10}}$$

$$\text{where}\quad \upsilon = \frac{\left(\frac{0.100138}{10}+\frac{0.0614622}{10}\right)^2}{\frac{\left(\frac{0.100138}{10}\right)^2}{10-1}+\frac{\left(\frac{0.0614622}{10}\right)^2}{10-1}} = 17.024 \cong 17$$

$$-1.426 \le (\mu_1-\mu_2) \le -0.889$$

2.45. Construct a data set for which the paired *t*-test statistic is very large, but for which the usual two-sample or pooled *t*-test statistic is small. In general, describe how you created the data. Does this give you any insight regarding how the paired *t*-test works?

A	B	delta
7.1662	8.2416	-1.0754
2.3590	2.4555	-0.0965
19.9977	21.1018	-1.1041
0.9077	2.3401	-1.4324
-15.9034	-15.0013	-0.9021
-6.0722	-5.5941	-0.4781
9.9501	10.6910	-0.7409
-1.0944	-0.1358	-0.9586
-4.6907	-3.3446	-1.3461
-6.6929	-5.9303	-0.7626

Minitab Output

Paired T-Test and Confidence Interval

```
Paired T for A - B
                  N      Mean     StDev   SE Mean
A                10      0.59     10.06      3.18
B                10      1.48     10.11      3.20
Difference       10    -0.890     0.398     0.126

95% CI for mean difference: (-1.174, -0.605)
T-Test of mean difference = 0 (vs not = 0): T-Value = -7.07  P-Value = 0.000
```

Two Sample T-Test and Confidence Interval

```
Two-sample T for A vs B

    N      Mean     StDev   SE Mean
A  10       0.6      10.1       3.2
B  10       1.5      10.1       3.2

Difference = mu A - mu B
Estimate for difference:  -0.89
95% CI for difference: (-10.37, 8.59)
T-Test of difference = 0 (vs not =): T-Value = -0.20  P-Value = 0.846  DF = 18
Both use Pooled StDev = 10.1
```

These two sets of data were created by making the observation for *A* and *B* moderately different within each pair (or block), but making the observations between pairs very different. The fact that the difference between pairs is large makes the pooled estimate of the standard deviation large and the two-sample *t*-test statistic small. Therefore the fairly small difference between the means of the two treatments that is present when they are applied to the same experimental unit cannot be detected. Generally, if the blocks are very different, then this will occur. Blocking eliminates the variability associated with the nuisance variable that they represent.

2.46. Consider the experiment described in problem 2.26. If the mean burning times of the two flames differ by as much as 2 minutes, find the power of the test. What sample size would be required to detect an actual difference in mean burning time of 1 minute with a power of at least 0.90?

From the *Minitab* output below, the power is 0.0740. This answer was obtained by using the pooled estimate of σ from Problem 2-11, S_p = 9.32. Because the difference in means is very small relative to the standard deviation, the power is very low.

Minitab Output

Power and Sample Size

```
2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Sigma = 9.32

            Sample
Difference    Size   Power
         2      10  0.0740
```

From the *Minitab* output below, the required sample size is 1827. The sample size is huge because the difference in means is very small relative to the standard deviation.

Minitab Output

Power and Sample Size

```
2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Sigma = 9.32

            Sample  Target  Actual
Difference    Size   Power   Power
         1    1827  0.9000  0.9001
```

2.47. Reconsider the bottle filling experiment described in Problem 2.24. Rework this problem assuming that the two population variances are unknown but equal.

Minitab Output

Two-Sample T-Test and CI: Machine 1, Machine 2

```
Two-sample T for Machine 1 vs Machine 2

          N      Mean     StDev   SE Mean
Machine  10   16.0150    0.0303    0.0096
Machine  10   16.0050    0.0255    0.0081

Difference = mu Machine 1 - mu Machine 2
Estimate for difference:  0.0100
95% CI for difference: (-0.0163, 0.0363)
T-Test of difference = 0 (vs not =): T-Value = 0.80  P-Value = 0.435  DF = 18
Both use Pooled StDev = 0.0280
```

The hypothesis test is the same: H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$

The conclusions are the same as Problem 2.19, do not reject H_0. There is no difference in the machines. The *P*-value for this analysis is 0.435.

The confidence interval is (-0.0163, 0.0363). This interval contains 0. There is no difference in machines.

2.48. Consider the data from problem 2.24. If the mean fill volume of the two machines differ by as much as 0.25 ounces, what is the power of the test used in problem 2.19? What sample size could result in a power of at least 0.9 if the actual difference in mean fill volume is 0.25 ounces?

The power is 1.0000 as shown in the analysis below.

Minitab Output

Power and Sample Size

```
2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Sigma = 0.028

            Sample
Difference    Size   Power
      0.25      10  1.0000
```

The required sample size is 2 as shown below.

Minitab Output

Power and Sample Size

```
2-Sample t Test

Testing mean 1 = mean 2 (versus not =)
Calculating power for mean 1 = mean 2 + difference
Alpha = 0.05  Sigma = 0.028

            Sample  Target  Actual
Difference    Size   Power   Power
      0.25       2  0.9000  0.9805
```

Chapter 3
Experiments with a Single Factor: The Analysis of Variance
Solutions

3.1. An experimenter has conducted a single-factor experiment with four levels of the factor, and each factor level has been replicated six times. The computed value of the *F*-statistic is $F_0 = 3.26$. Find bounds on the *P*-value.

Table *P*-value = 0.025, 0.050 Computer *P*-value = 0.043

3.2. An experimenter has conducted a single-factor experiment with six levels of the factor, and each factor level has been replicated three times. The computed value of the *F*-statistic is $F_0 = 5.81$. Find bounds on the *P*-value.

Table *P*-value < 0.010 Computer *P*-value = 0.006

3.3. A computer ANOVA output is shown below. Fill in the blanks. You may give bounds on the *P*-value.

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	3	36.15	?	?	?
Error	?	?	?		
Total	19	196.04			

Completed table is:

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	3	36.15	12.05	1.21	0.3395
Error	16	159.89	9.99		
Total	19	196.04			

3.4. A computer ANOVA output is shown below. Fill in the blanks. You may give bounds on the *P*-value.

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	?	?	246.93	?	?
Error	25	186.53	?		
Total	29	1174.24			

Completed table is:

One-way ANOVA

Source	DF	SS	MS	F	P
Factor	4	987.71	246.93	33.09	< 0.0001
Error	25	186.53	7.46		
Total	29	1174.24			

3.5. An article appeared in *The Wall Street Journal* on Tuesday, April 27, 2010, with the title "Eating Chocolate Is Linked to Depression." The article reported on a study funded by the National Heart, Lung and Blood Institute (part of the National Institutes of Health) and conducted by the faculty at the University of California, San Diego, and the University of California, Davis. The research was also published in the *Archives of Internal Medicine* (2010, pp. 699-703). The study examined 931 adults who were not taking antidepressants and did not have known cardiovascular disease or diabetes. The group was about 70% men and the average age of the group was reported to be about 58. The participants were asked about chocolate consumption and then screened for depression using a questionnaire. People who scored less than 16 on the questionnaire are not considered depressed, while those with scores above 16 and less than or equal to 22 are considered possibly depressed, while those with scores above 22 are considered likely to be depressed. The survey found that people who were not depressed ate an average of 8.4 servings of chocolate per month, while those individuals who scored above 22 were likely to be depressed ate the most chocolate, an average of 11.8 servings per month. No differentiation was made between dark and milk chocolate. Other foods were also examined, but no patterned emerged between other foods and depression. Is this study really a designed experiment? Does it establish a cause-and-effect link between chocolate consumption and depression? How would the study have to be conducted to establish such a link?

This is not a designed experiment, and it does not establish a cause-and-effect link between chocolate consumption and depression. An experiment could be run by giving a group of people a defined amount of chocolate servings per month for several months, while not giving another group any chocolate. Ideally it would be good to have the participants not eat any chocolate for a period of time before the experiment, and measure depression for each participant before and after the experiment.

3.6. An article in *Bioelectromagnetics* ("Electromagnetic Effects on Forearm Disuse Osteopenia: A Randomized, Double-Blind, Sham-Controlled Study," Vol. 32, 2011, pp. 273 – 282) describes a randomized, double-blind, sham-controlled, feasibility and dosing study to determine if a common pulsing electromagnetic field (PEMF) treatment could moderate the substantial osteopenia that occurs after forearm disuse. Subjects were randomized into four groups after a distal radius fracture, or carpal surgery requiring immobilization in a cast. Active of identical sham PEMF transducers were worn on a distal forearm for 1, 2, or 4h/day for 8 weeks starting after cast removal ("baseline") when bone density continues to decline. Bone mineral density (BMD) and bone geometry were measured in the distal forearm by dual energy X-ray absorptiometry (DXA) and peripheral quantitative computed tomography (pQCT). The data below are the percent losses in BMD measurements on the radius after 16weeks for patients wearing the active or sham PEMF transducers for 1, 2, or 4h/day (data were constructed to match the means and standard deviations read from a graph in the paper).

Sham	PEMF 1h/day	PEMF 2h/day	PEMF 4h/day
4.51	5.32	4.73	7.03
7.95	6.00	5.81	4.65
4.97	5.12	5.69	6.65

3.00	7.08	3.86	5.49
7.97	5.48	4.06	6.98
2.23	6.52	6.56	4.85
3.95	4.09	8.34	7.26
5.64	6.28	3.01	5.92
9.35	7.77	6.71	5.58
6.52	5.68	6.51	7.91
4.96	8.47	1.70	4.90
6.10	4.58	5.89	4.54
7.19	4.11	6.55	8.18
4.03	5.72	5.34	5.42
2.72	5.91	5.88	6.03
9.19	6.89	7.50	7.04
5.17	6.99	3.28	5.17
5.70	4.98	5.38	7.60
5.85	9.94	7.30	7.90
6.45	6.38	5.46	7.91

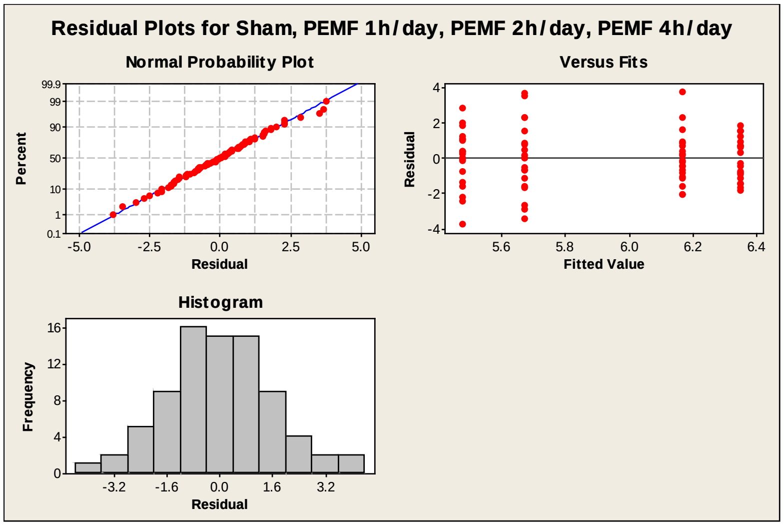
(a) Is there evidence to support a claim that PEMF usage affects BMD loss? If so, analyze the data to determine which specific treatments produce the differences. The ANOVA from the Minitab output shows that there is no difference between the treatments; P=0.281.

Minitab Output

One-way ANOVA: Sham, PEMF 1h/day, PEMF 2h/day, PEMF 4h/day

```
Source  DF      SS    MS     F      P
Factor   3   10.04  3.35  1.30  0.281
Error   76  196.03  2.58
Total   79  206.07

S = 1.606   R-Sq = 4.87%   R-Sq(adj) = 1.12%


                               Individual 95% CIs For Mean Based on
                               Pooled StDev
Level         N   Mean  StDev   -+---------+---------+---------+--------
Sham         20  5.673  2.002       (-----------*----------)
PEMF 1h/day  20  6.165  1.444               (-----------*-----------)
PEMF 2h/day  20  5.478  1.645   (-----------*-----------)
PEMF 4h/day  20  6.351  1.232                  (-----------*-----------)
                                -+---------+---------+---------+--------
                               4.80      5.40      6.00      6.60
```

(b) Analyze the residuals from this experiment and comment on the underlying assumptions and model adequacy. The residuals show the model is good.



3.7. The tensile strength of Portland cement is being studied. Four different mixing techniques can be used economically. A completely randomized experiment was conducted and the following data were collected.

Mixing Technique	Tensile Strength (lb/in^2)			
1	3129	3000	2865	2890
2	3200	3300	2975	3150
3	2800	2900	2985	3050
4	2600	2700	2600	2765

(a) Test the hypothesis that mixing techniques affect the strength of the cement. Use $\alpha = 0.05$.

Design Expert Output

Response: Tensile Strengthin lb/in^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	4.897E+005	3	1.632E+005	12.73	0.0005	significant
A	*4.897E+005*	*3*	*1.632E+005*	*12.73*	*0.0005*	
Residual	1.539E+005	12	12825.69			
Lack of Fit	*0.000*	*0*				
Pure Error	*1.539E+005*	*12*	*12825.69*			
Cor Total	6.436E+005	15				

The Model F-value of 12.73 implies the model is significant. There is only a 0.05% chance that a "Model F-Value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	2971.00	56.63
2-2	3156.25	56.63
3-3	2933.75	56.63
4-4	2666.25	56.63

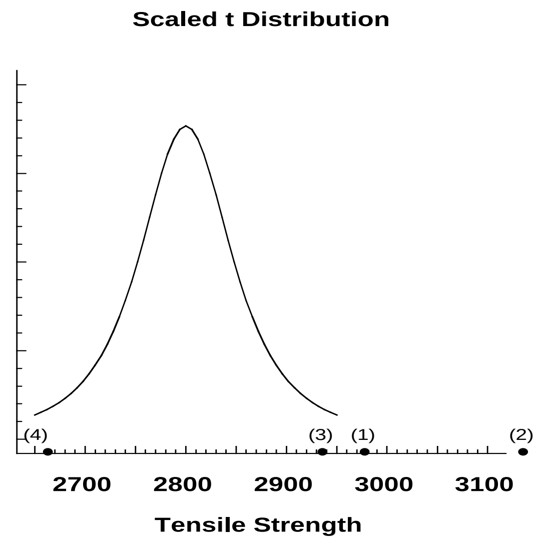
Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	-185.25	1	80.08	-2.31	0.0392
1 vs 3	37.25	1	80.08	0.47	0.6501
1 vs 4	304.75	1	80.08	3.81	0.0025

2 vs 3	222.50	1	80.08	2.78	0.0167
2 vs 4	490.00	1	80.08	6.12	< 0.0001
3 vs 4	267.50	1	80.08	3.34	0.0059

The *F*-value is 12.73 with a corresponding *P*-value of .0005. Mixing technique has an effect.

(b) Construct a graphical display as described in Section 3.5.3 to compare the mean tensile strengths for the four mixing techniques. What are your conclusions?

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{n}} = \sqrt{\frac{12825.7}{4}} = 56.625$$



Based on examination of the plot, we would conclude that μ_1 and μ_3 are the same; that μ_4 differs from μ_1 and μ_3, that μ_2 differs from μ_1 and μ_3, and that μ_2 and μ_4 are different.

(c) Use the Fisher LSD method with α=0.05 to make comparisons between pairs of means.

$$LSD = t_{\alpha/2, N-a}\sqrt{\frac{2MS_E}{n}}$$

$$LSD = t_{0.025,16-4}\sqrt{\frac{2(12825.7)}{4}}$$

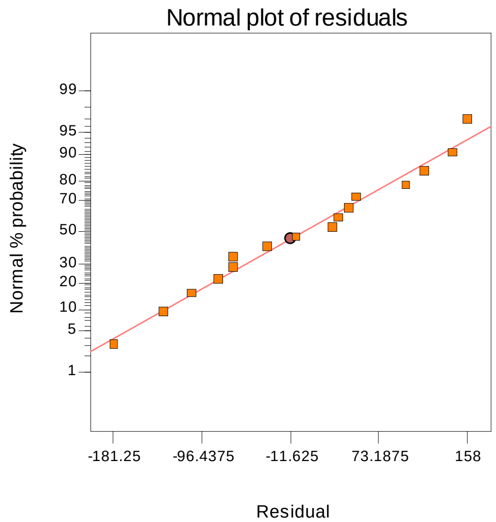
$$LSD = 2.179\sqrt{6412.85} = 174.495$$

Treatment 2 vs. Treatment 4 = 3156.250 - 2666.250 = 490.000 > 174.495
Treatment 2 vs. Treatment 3 = 3156.250 - 2933.750 = 222.500 > 174.495
Treatment 2 vs. Treatment 1 = 3156.250 - 2971.000 = 185.250 > 174.495
Treatment 1 vs. Treatment 4 = 2971.000 - 2666.250 = 304.750 > 174.495
Treatment 1 vs. Treatment 3 = 2971.000 - 2933.750 = 37.250 < 174.495
Treatment 3 vs. Treatment 4 = 2933.750 - 2666.250 = 267.500 > 174.495

The Fisher LSD method is also presented in the Design-Expert computer output above. The results agree with the graphical method for this experiment.

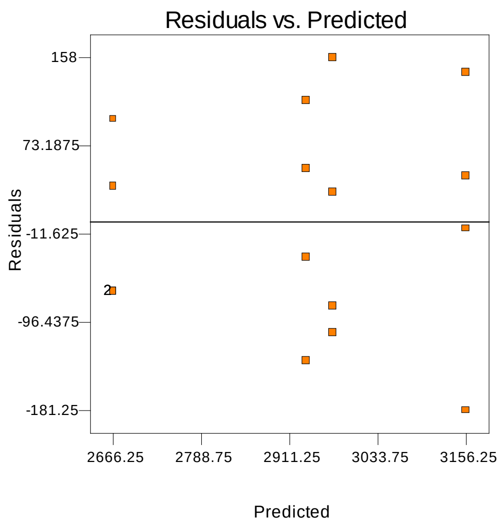
(d) Construct a normal probability plot of the residuals. What conclusion would you draw about the validity of the normality assumption?

There is nothing unusual about the normal probability plot of residuals.



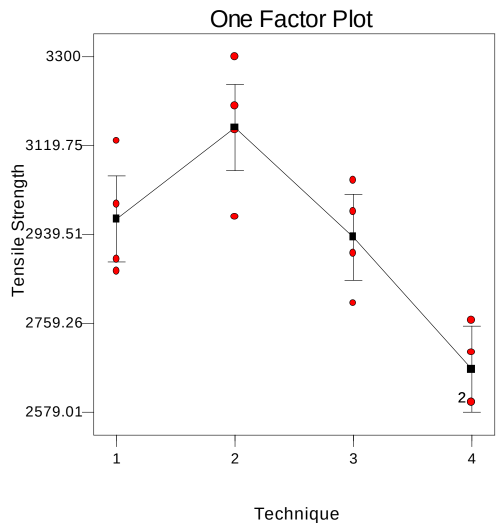
(e) Plot the residuals versus the predicted tensile strength. Comment on the plot.

There is nothing unusual about this plot.



(f) Prepare a scatter plot of the results to aid the interpretation of the results of this experiment.

Design-Expert automatically generates the scatter plot. The plot below also shows the sample average for each treatment and the 95 percent confidence interval on the treatment mean.



3.8. (a) Rework part (c) of Problem 3.7 using Tukey's test with $\alpha = 0.05$. Do you get the same conclusions from Tukey's test that you did from the graphical procedure and/or the Fisher LSD method?

Minitab Output

```
Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.0117
Critical value = 4.20

Intervals for (column level mean) - (row level mean)

             1        2        3

    2     -423
            53

    3     -201      -15
           275      460

    4       67      252       30
           543      728      505
```

No, the conclusions are not the same. The mean of Treatment 4 is different than the means of Treatments 1, 2, and 3. However, the mean of Treatment 2 is not different from the means of Treatments 1 and 3 according to Tukey's method, they were found to be different using the graphical method and the Fisher LSD method.

(b) Explain the difference between the Tukey and Fisher procedures.

Both Tukey and Fisher utilize a single critical value; however, Tukey's is based on the studentized range statistic while Fisher's is based on *t* distribution.

3.9. Reconsider the experiment in Problem 3.7. Find a 95 percent confidence interval on the mean tensile strength of the portland cement produced by each of the four mixing techniques. Also find a 95 percent confidence interval on the difference in means for techniques 1 and 3. Does this aid in interpreting the results of the experiment?

$$\bar{y}_{i.} - t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}} \le \mu_i \le \bar{y}_{i.} + t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}}$$

$$\text{Treatment 1: } 2971 \pm 2.179\sqrt{\frac{12825.69}{4}}$$

$$2971 \pm 123.387$$

$$2847.613 \le \mu_1 \le 3094.387$$

$$\text{Treatment 2: } 3156.25 \pm 123.387$$

$$3032.863 \le \mu_2 \le 3279.637$$

$$\text{Treatment 3: } 2933.75 \pm 123.387$$

$$2810.363 \le \mu_3 \le 3057.137$$

$$\text{Treatment 4: } 2666.25 \pm 123.387$$

$$2542.863 \le \mu_4 \le 2789.637$$

$$\text{Treatment 1 - Treatment 3: } \bar{y}_{i.} - \bar{y}_{j.} - t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}} \le \mu_i - \mu_j \le \bar{y}_{i.} - \bar{y}_{j.} + t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}}$$

$$2971.00 - 2933.75 \pm 2.179\sqrt{\frac{2(12825.7)}{4}}$$

$$-137.245 \le \mu_1 - \mu_3 \le 211.745$$

Because the confidence interval for the difference between means 1 and 3 spans zero, we agree with the statement in Problem 3.5 (b); there is not a statistical difference between these two means.

3.10. A product developer is investigating the tensile strength of a new synthetic fiber that will be used to make cloth for men's shirts. Strength is usually affected by the percentage of cotton used in the blend of materials for the fiber. The engineer conducts a completely randomized experiment with five levels of cotton content and replicated the experiment five times. The data are shown in the following table.

Cotton Weight Percentage			Observations		
15	7	7	15	11	9
20	12	17	12	18	18
25	14	19	19	18	18
30	19	25	22	19	23
35	7	10	11	15	11

(a) Is there evidence to support the claim that cotton content affects the mean tensile strength? Use α = 0.05.

Minitab Output

One-way ANOVA: Tensile Strength versus Cotton Percentage

```
Analysis of Variance for Tensile
Source     DF        SS        MS        F        P
Cotton P    4    475.76    118.94    14.76    0.000
Error      20    161.20      8.06
Total      24    636.96
```

Yes, the *F*-value is 14.76 with a corresponding *P*-value of 0.000. The percentage of cotton in the fiber appears to have an affect on the tensile strength.

(b) Use the Fisher LSD method to make comparisons between the pairs of means. What conclusions can you draw?

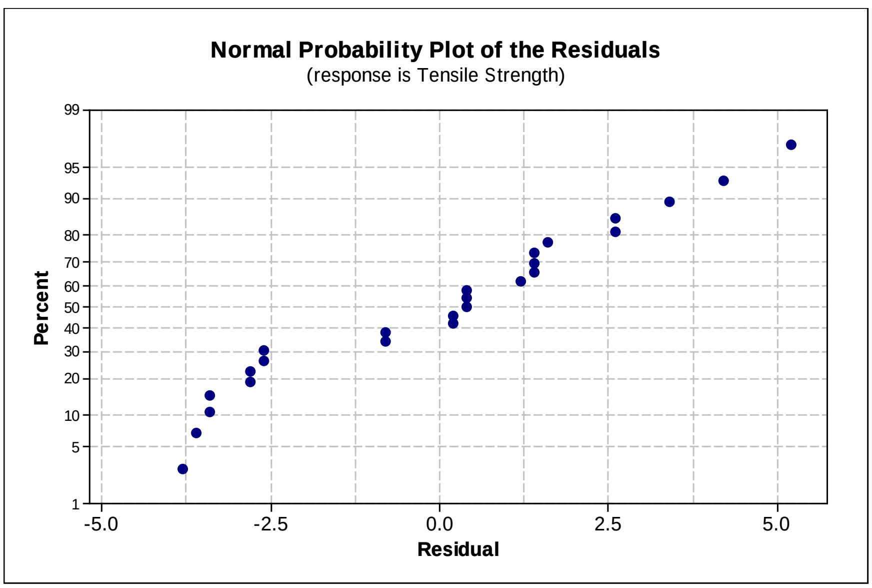
Minitab Output

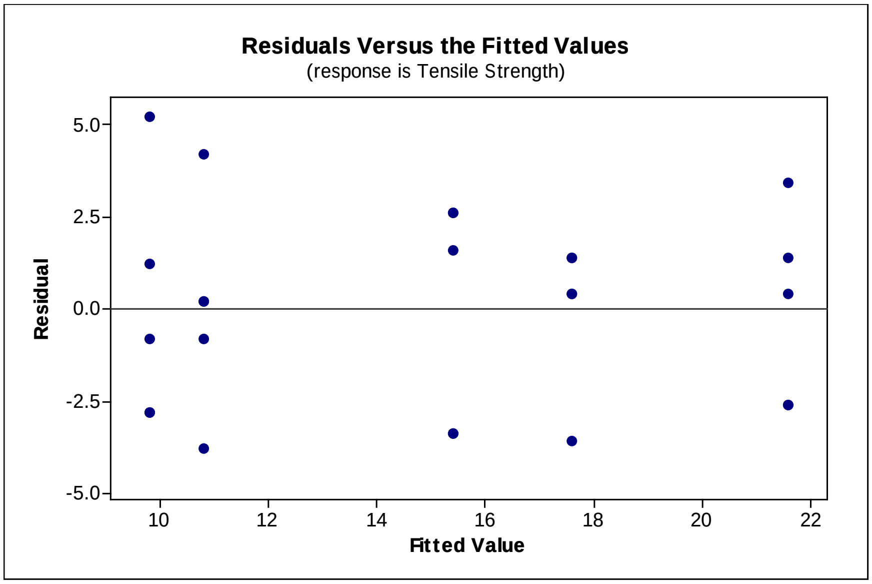
```
Fisher's pairwise comparisons

    Family error rate = 0.264
Individual error rate = 0.0500

Critical value = 2.086

Intervals for (column level mean) - (row level mean)

                  15          20          25          30

      20      -9.346
              -1.854

      25     -11.546      -5.946
              -4.054       1.546

      30     -15.546      -9.946      -7.746
              -8.054      -2.454      -0.254

      35      -4.746       0.854       3.054       7.054
               2.746       8.346      10.546      14.546
```

In the Minitab output the pairs of treatments that do not contain zero in the pair of numbers indicates that there is a difference in the pairs of the treatments. 15% cotton is different than 20%, 25% and 30%. 20% cotton is different than 30% and 35% cotton. 25% cotton is different than 30% and 35% cotton. 30% cotton is different than 35%.

(c) Analyze the residuals from this experiment and comment on model adequacy.

The residual plots below show nothing unusual.





3.11. Reconsider the experiment described in Problem 3.10. Suppose that 30 percent cotton content is a control. Use Dunnett's test with $\alpha = 0.05$ to compare all of the other means with the control.

For this problem: $a = 5$, $a\text{-}1 = 4$, f=20, n=5 and $\alpha = 0.05$

$$d_{0.05}(4,20)\sqrt{\frac{2MS_E}{n}} = 2.65\sqrt{\frac{2(8.06)}{5}} = 4.76$$

$$\begin{aligned}
&1 \text{ vs. } 4: \bar{y}_{1.} - \bar{y}_{4.} = 9.8 - 21.6 = -11.8^* \\
&2 \text{ vs. } 4: \bar{y}_{2.} - \bar{y}_{4.} = 15.4 - 21.6 = -6.2^* \\
&3 \text{ vs. } 4: \bar{y}_{3.} - \bar{y}_{4.} = 17.6 - 21.6 = -4.0 \\
&5 \text{ vs. } 4: \bar{y}_{5.} - \bar{y}_{4.} = 10.8 - 21.6 = -10.8^*
\end{aligned}$$

The control treatment, treatment 4, differs from treatments 1, 2 and 5.

3.12. A pharmaceutical manufacturer wants to investigate the bioactivity of a new drug. A completely randomized single-factor experiment was conducted with three dosage levels, and the following results were obtained.

Dosage	Observations			
20g	24	28	37	30
30g	37	44	31	35
40g	42	47	52	38

(a) Is there evidence to indicate that dosage level affects bioactivity? Use $\alpha = 0.05$.

Minitab Output

```
One-way ANOVA: Activity versus Dosage

Analysis of Variance for Activity
Source     DF        SS        MS        F        P
Dosage      2     450.7     225.3     7.04    0.014
Error       9     288.3      32.0
Total      11     738.9
```

There appears to be a different in the dosages.

(b) If it is appropriate to do so, make comparisons between the pairs of means. What conclusions can you draw?

Because there appears to be a difference in the dosages, the comparison of means is appropriate.

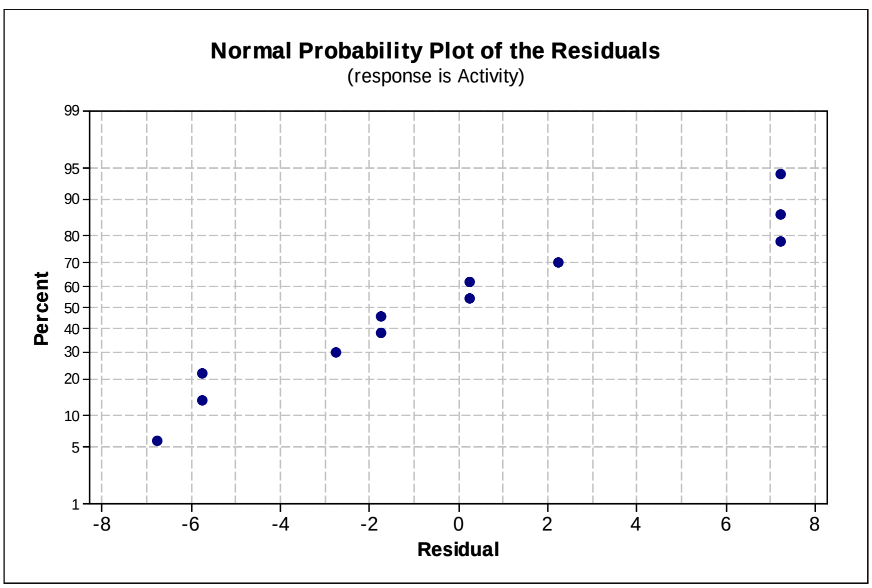
Minitab Output

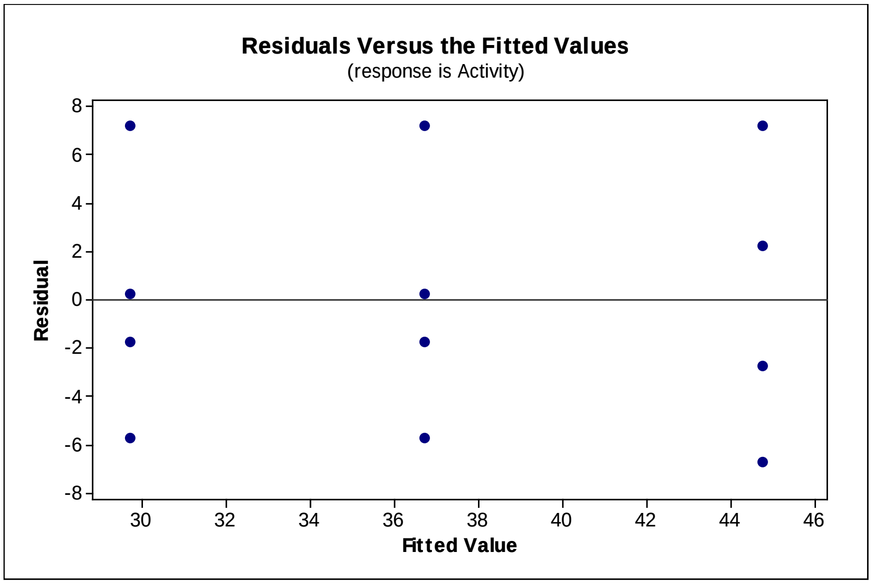
```
Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.0209

Critical value = 3.95

Intervals for (column level mean) - (row level mean)

               20g         30g

     30g     -18.177
               4.177

     40g     -26.177     -19.177
              -3.823       3.177
```

The Tukey comparison shows a difference in the means between the 20g and the 40g dosages.

(c) Analyze the residuals from this experiment and comment on the model adequacy.

There is nothing too unusual about the residual plots shown below.





3.13. A rental car company wants to investigate whether the type of car rented affects the length of the rental period. An experiment is run for one week at a particular location, and 10 rental contracts are selected at random for each car type. The results are shown in the following table.

Type of Car	Observations									
Sub-compact	3	5	3	7	6	5	3	2	1	6
Compact	1	3	4	7	5	6	3	2	1	7
Midsize	4	1	3	5	7	1	2	4	2	7
Full Size	3	5	7	5	10	3	4	7	2	7

(a) Is there evidence to support a claim that the type of car rented affects the length of the rental contract? Use $\alpha = 0.05$. If so, which types of cars are responsible for the difference?

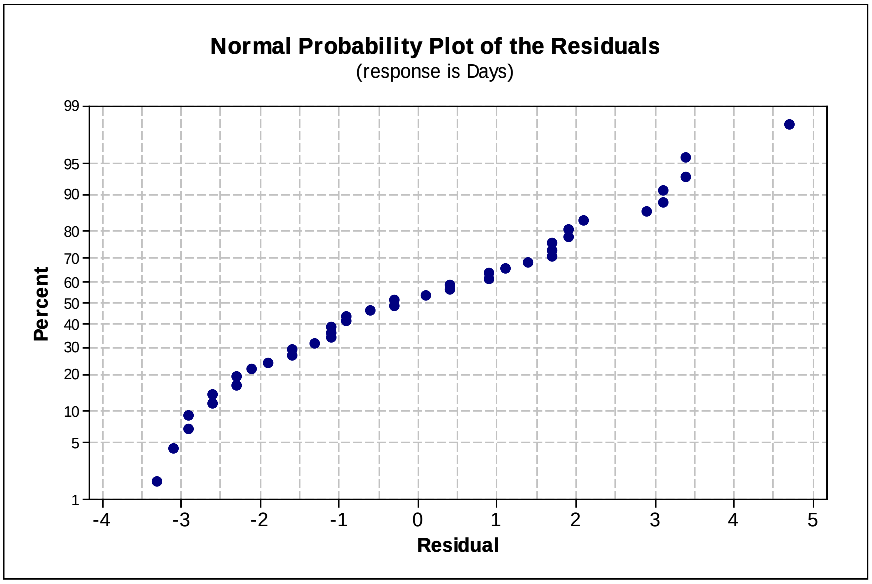
Minitab Output

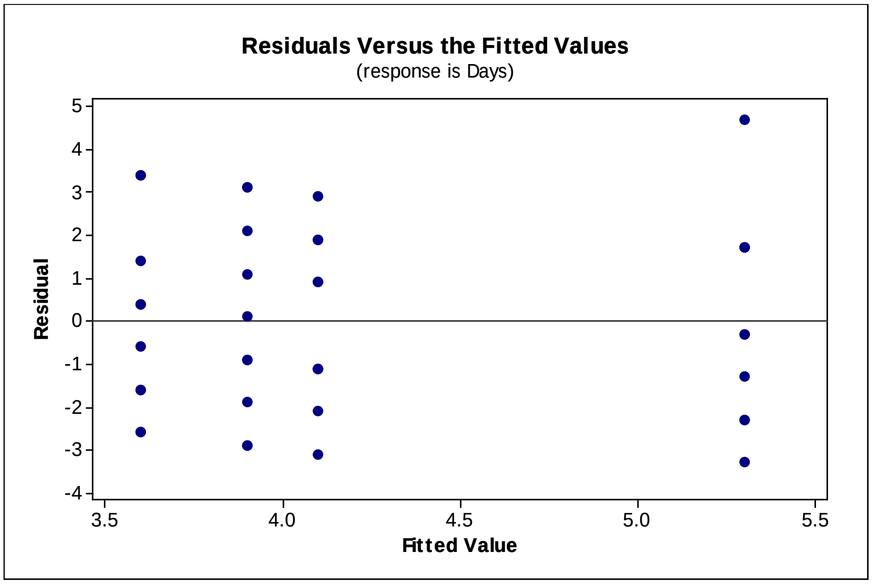
One-way ANOVA: Days versus Car Type

```
Analysis of Variance for Days
Source      DF        SS        MS        F        P
Car Type     3     16.68      5.56     1.11    0.358
Error       36    180.30      5.01
Total       39    196.98
```

There is no difference.

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





There is nothing unusual about the residuals.

(c) Notice that the response variable in this experiment is a count. Should this cause any potential concerns about the validity of the analysis of variance?

Because the data is count data, a square root transformation could be applied. The analysis is shown below. It does not change the interpretation of the data.

Minitab Output

One-way ANOVA: Sqrt Days versus Car Type

```
Analysis of Variance for Sqrt Day
Source      DF        SS        MS        F        P
Car Type     3     1.087     0.362     1.10    0.360
Error       36    11.807     0.328
Total       39    12.893
```

3.14. I belong to a golf club in my neighborhood. I divide the year into three golf seasons: summer (June-September), winter (November-March) and shoulder (October, April and May). I believe that I play my best golf during the summer (because I have more time and the course isn't crowded) and shoulder (because the course isn't crowded) seasons, and my worst golf during the winter (because all of the part-year residents show up, and the course is crowded, play is slow, and I get frustrated). Data from the last year are shown in the following table.

Season	Observations								
Summer	83	85	85	87	90	88	88	84	91	90
Shoulder	91	87	84	87	85	86	83			
Winter	94	91	87	85	87	91	92	86		

(a) Do the data indicate that my opinion is correct? Use $\alpha = 0.05$.

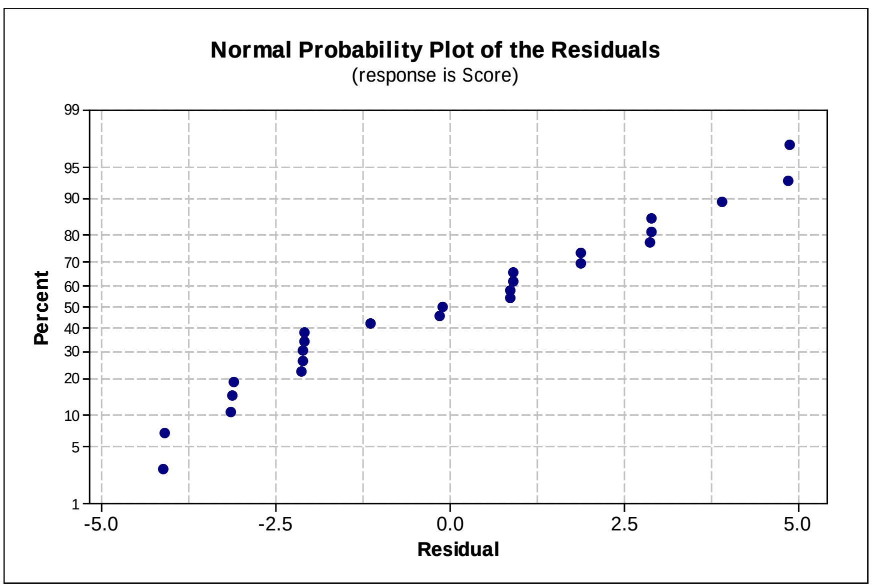
Minitab Output

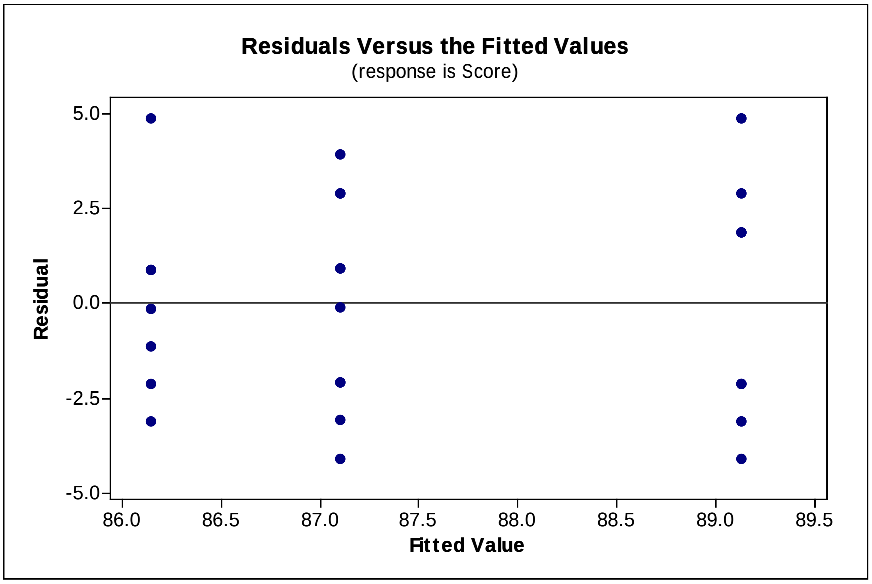
One-way ANOVA: Score versus Season

```
Analysis of Variance for Score
Source     DF        SS        MS        F        P
Season      2     35.61     17.80     2.12    0.144
Error      22    184.63      8.39
Total      24    220.24
```

The data do not support the author's opinion.

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





There is nothing unusual about the residuals.

3.15. A regional opera company has tried three approaches to solicit donations from 24 potential sponsors. The 24 potential sponsors were randomly divided into three groups of eight, and one approach was used for each group. The dollar amounts of the resulting contributions are shown in the following table.

Approach	Contributions (in $)							
1	1000	1500	1200	1800	1600	1100	1000	1250
2	1500	1800	2000	1200	2000	1700	1800	1900
3	900	1000	1200	1500	1200	1550	1000	1100

(a) Do the data indicate that there is a difference in results obtained from the three different approaches? Use $\alpha = 0.05$.

Minitab Output

One-way ANOVA: Contribution versus Approach

```
Analysis of Variance for Contribution
Source     DF        SS        MS        F        P
Approach    2   1362708    681354     9.41    0.001
Error      21   1520625     72411
Total      23   2883333
```

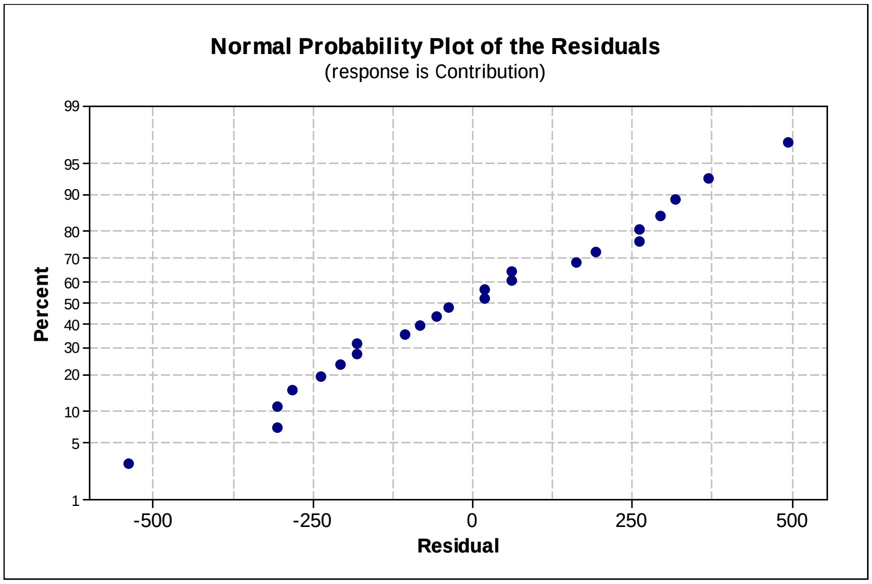
There is a difference between the approaches. The Tukey test will indicate which are different. Approach 2 is different than approach 1 and approach 3.

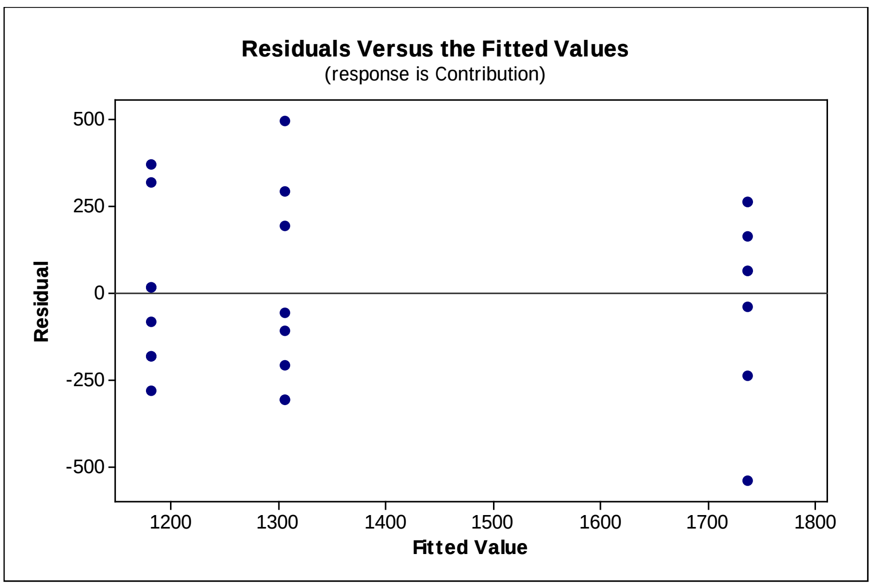
Minitab Output

```
Tukey's pairwise comparisons

    Family error rate = 0.0500
Individual error rate = 0.0200

Critical value = 3.56

Intervals for (column level mean) - (row level mean)

                 1           2
```

```
      2        -770
                -93

      3        -214        218
                464        895
```

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





There is nothing unusual about the residuals.

3.16. An experiment was run to determine whether four specific firing temperatures affect the density of a certain type of brick. A completely randomized experiment led to the following data:

Temperature	Density				
100	21.8	21.9	21.7	21.6	21.7
125	21.7	21.4	21.5	21.4	
150	21.9	21.8	21.8	21.6	21.5
175	21.9	21.7	21.8	21.4	

(a) Does the firing temperature affect the density of the bricks? Use $\alpha = 0.05$.

No, firing temperature does not affect the density of the bricks. Refer to the Design-Expert output below.

Design Expert Output

Response: Density
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	0.16	3	0.052	2.02	0.1569	not significant
A	*0.16*	*3*	*0.052*	*2.02*	*0.1569*	
Residual	0.36	14	0.026			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.36*	*14*	*0.026*			
Cor Total	0.52	17				

The "Model F-value" of 2.02 implies the model is not significant relative to the noise. There is a 15.69 % chance that a "Model F-value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

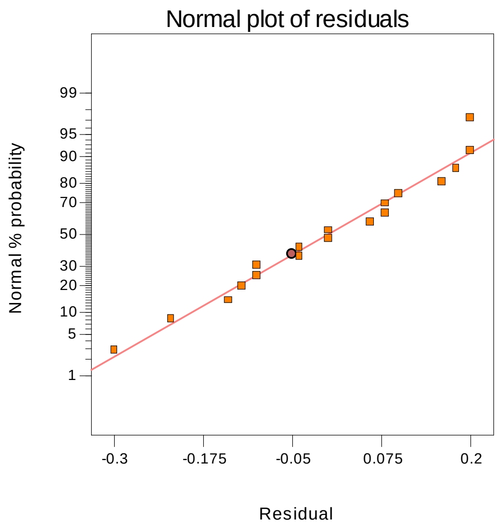
	Estimated Mean	Standard Error
1-100	21.74	0.072
2-125	21.50	0.080
3-150	21.72	0.072
4-175	21.70	0.080

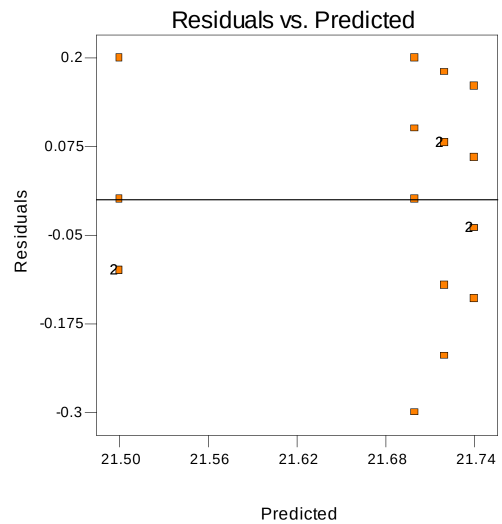
Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	0.24	1	0.11	2.23	0.0425
1 vs 3	0.020	1	0.10	0.20	0.8465
1 vs 4	0.040	1	0.11	0.37	0.7156
2 vs 3	-0.22	1	0.11	-2.05	0.0601
2 vs 4	-0.20	1	0.11	-1.76	0.0996
3 vs 4	0.020	1	0.11	0.19	0.8552

(b) Is it appropriate to compare the means using the Fisher LSD method in this experiment?

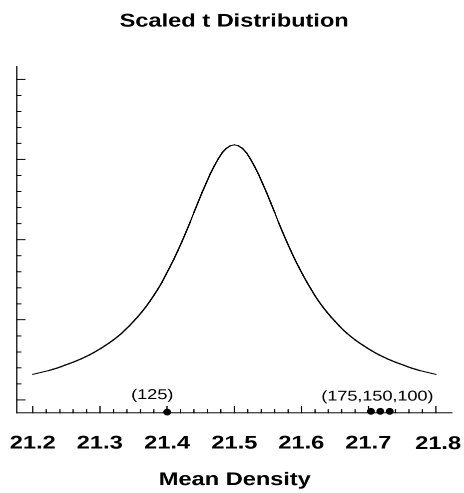
The analysis of variance tells us that there is no difference in the treatments. There is no need to proceed with Fisher's LSD method to decide which mean is difference.

(c) Analyze the residuals from this experiment. Are the analysis of variance assumptions satisfied? There is nothing unusual about the residual plots.





(d) Construct a graphical display of the treatments as described in Section 3.5.3. Does this graph adequately summarize the results of the analysis of variance in part (b). Yes.



3.17. Rework Part (d) of Problem 3.16 using the Tukey method. What conclusions can you draw? Explain carefully how you modified the procedure to account for unequal sample sizes.

When sample sizes are unequal, the appropriate formula for the Tukey method is

$$T_\alpha = \frac{q_\alpha(a,f)}{\sqrt{2}}\sqrt{MS_E\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$$

$$\begin{aligned}
&\text{Treatment 1 vs. Treatment 2} = 21.74 - 21.50 = 0.24 < 0.314\\
&\text{Treatment 1 vs. Treatment 3} = 21.74 - 21.72 = 0.02 < 0.296\\
&\text{Treatment 1 vs. Treatment 4} = 21.74 - 21.70 = 0.04 < 0.314\\
&\text{Treatment 3 vs. Treatment 2} = 21.72 - 21.50 = 0.22 < 0.314\\
&\text{Treatment 4 vs. Treatment 2} = 21.70 - 21.50 = 0.20 < 0.331\\
&\text{Treatment 3 vs. Treatment 4} = 21.72 - 21.70 = 0.02 < 0.314
\end{aligned}$$

All pairwise comparisons do not identify differences. Notice that there are different critical values for the comparisons depending on the sample sizes of the two groups being compared.

Because we could not reject the hypothesis of equal means using the analysis of variance, we should never have performed the Tukey test (or any other multiple comparison procedure, for that matter). If you ignore the analysis of variance results and run multiple comparisons, you will likely make type I errors.

3.18. A manufacturer of television sets is interested in the effect of tube conductivity of four different types of coating for color picture tubes. A completely randomized experiment is conducted and the following conductivity data are obtained:

Coating Type	Conductivity			
1	143	141	150	146
2	152	149	137	143
3	134	136	132	127
4	129	127	132	129

(a) Is there a difference in conductivity due to coating type? Use $\alpha = 0.05$.

Yes, there is a difference in means. Refer to the Design-Expert output below..

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	844.69	3	281.56	14.30	0.0003	significant
A	*844.69*	*3*	*281.56*	*14.30*	*0.0003*	
Residual	236.25	12	19.69			
Lack of Fit	*0.000*	*0*				
Pure Error	*236.25*	*12*	*19.69*			
Cor Total	1080.94	15				

The Model F-value of 14.30 implies the model is significant. There is only a 0.03% chance that a "Model F-Value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	145.00	2.22
2-2	145.25	2.22
3-3	132.25	2.22
4-4	129.25	2.22

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-0.25	1	3.14	-0.080	0.9378
1 vs 3	12.75	1	3.14	4.06	0.0016
1 vs 4	15.75	1	3.14	5.02	0.0003
2 vs 3	13.00	1	3.14	4.14	0.0014
2 vs 4	16.00	1	3.14	5.10	0.0003
3 vs 4	3.00	1	3.14	0.96	0.3578

(b) Estimate the overall mean and the treatment effects.

$$\hat{\mu} = 2207/16 = 137.9375$$
$$\hat{\tau}_1 = \bar{y}_{1.} - \bar{y}_{..} = 145.00 - 137.9375 = 7.0625$$
$$\hat{\tau}_2 = \bar{y}_{2.} - \bar{y}_{..} = 145.25 - 137.9375 = 7.3125$$
$$\hat{\tau}_3 = \bar{y}_{3.} - \bar{y}_{..} = 132.25 - 137.9375 = -5.6875$$
$$\hat{\tau}_4 = \bar{y}_{4.} - \bar{y}_{..} = 129.25 - 137.9375 = -8.6875$$

(c) Compute a 95 percent interval estimate of the mean of coating type 4. Compute a 99 percent interval estimate of the mean difference between coating types 1 and 4.

$$\text{Treatment 4: } 129.25 \pm 2.179\sqrt{\frac{19.69}{4}}$$
$$124.4155 \le \mu_4 \le 134.0845$$
$$\text{Treatment 1 - Treatment 4: } (145 - 129.25) \pm 3.055\sqrt{\frac{(2)19.69}{4}}$$
$$6.164 \le \mu_1 - \mu_4 \le 25.336$$

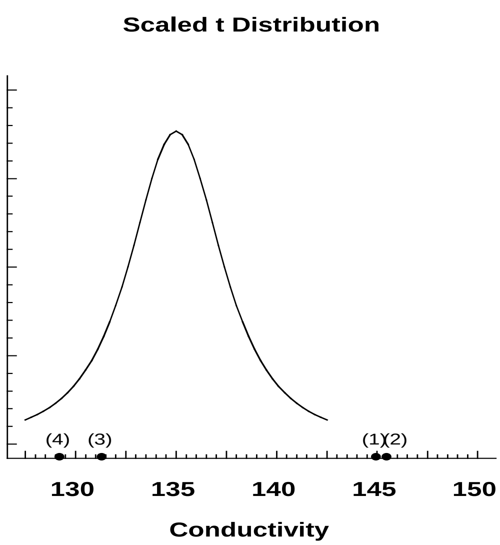
(d) Test all pairs of means using the Fisher LSD method with α=0.05.

Refer to the Design-Expert output above. The Fisher LSD procedure is automatically included in the output.

The means of Coating Type 2 and Coating Type 1 are not different. The means of Coating Type 3 and Coating Type 4 are not different. However, Coating Types 1 and 2 produce higher mean conductivity than does Coating Types 3 and 4.

(e) Use the graphical method discussed in Section 3.5.3 to compare the means. Which coating produces the highest conductivity?

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{n}} = \sqrt{\frac{19.96}{4}} = 2.219 \text{ Coating types 1 and 2 produce the highest conductivity.}$$

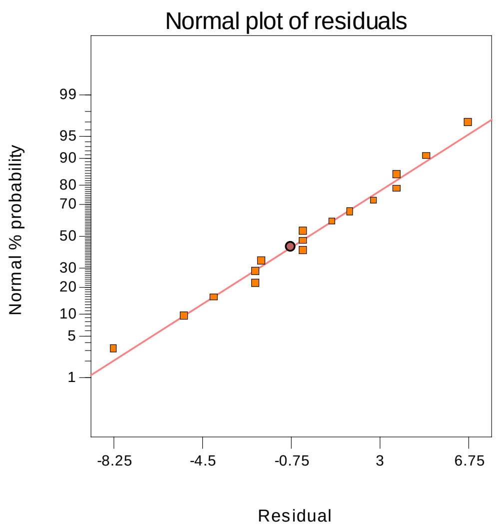


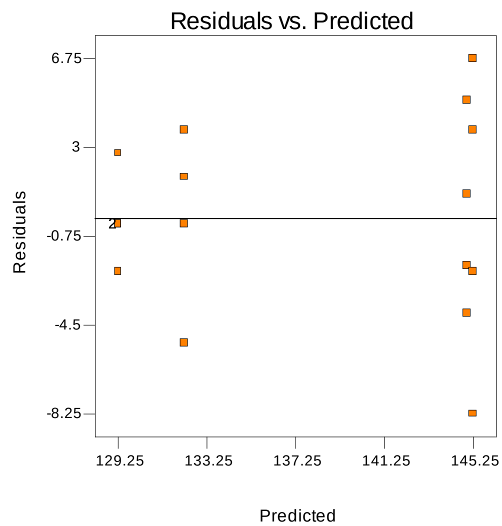
(f) Assuming that coating type 4 is currently in use, what are your recommendations to the manufacturer? We wish to minimize conductivity.

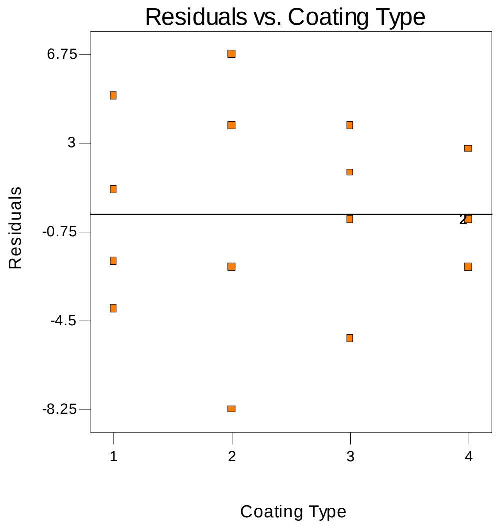
Since coatings 3 and 4 do not differ, and as they both produce the lowest mean values of conductivity, use either coating 3 or 4. As type 4 is currently being used, there is probably no need to change.

3.19. Reconsider the experiment in Problem 3.18. Analyze the residuals and draw conclusions about model adequacy.

There is nothing unusual in the normal probability plot. A funnel shape is seen in the plot of residuals versus predicted conductivity indicating a possible non-constant variance.







3.20. An article in the *ACI Materials Journal* (Vol. 84, 1987. pp. 213-216) describes several experiments investigating the rodding of concrete to remove entrapped air. A 3" x 6" cylinder was used, and the number of times this rod was used is the design variable. The resulting compressive strength of the concrete specimen is the response. The data are shown in the following table.

Rodding Level	Compressive Strength		
10	1530	1530	1440
15	1610	1650	1500
20	1560	1730	1530
25	1500	1490	1510

(a) Is there any difference in compressive strength due to the rodding level? Use $\alpha = 0.05$.

There are no differences.

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	28633.33	3	9544.44	1.87	0.2138	not significant
A	*28633.33*	*3*	*9544.44*	*1.87*	*0.2138*	
Residual	40933.33	8	5116.67			
Lack of Fit	*0.000*	*0*				
Pure Error	*40933.33*	*8*	*5116.67*			
Cor Total	69566.67	11				

The "Model F-value" of 1.87 implies the model is not significant relative to the noise. There is a 21.38 % chance that a "Model F-value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

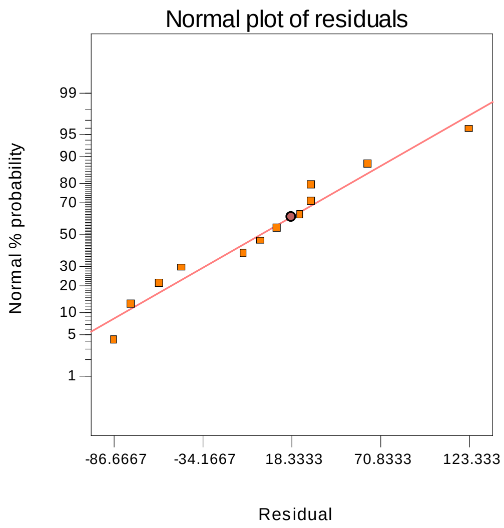
Estimated Mean		Standard Error
1-10	1500.00	41.30
2-15	1586.67	41.30
3-20	1606.67	41.30
4-25	1500.00	41.30

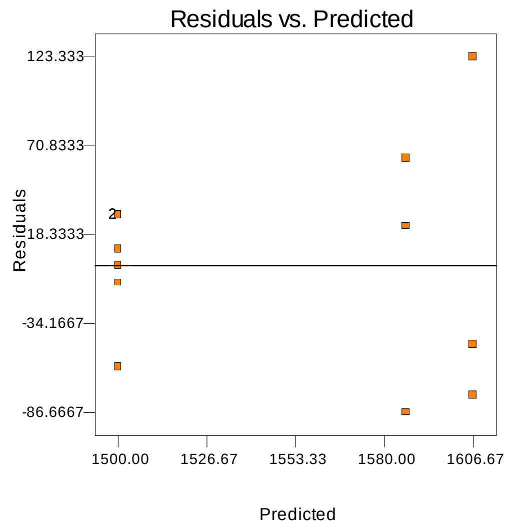
Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	-86.67	1	58.40	-1.48	0.1761
1 vs 3	-106.67	1	58.40	-1.83	0.1052
1 vs 4	0.000	1	58.40	0.000	1.0000
2 vs 3	-20.00	1	58.40	-0.34	0.7408
2 vs 4	86.67	1	58.40	1.48	0.1761
3 vs 4	106.67	1	58.40	1.83	0.1052

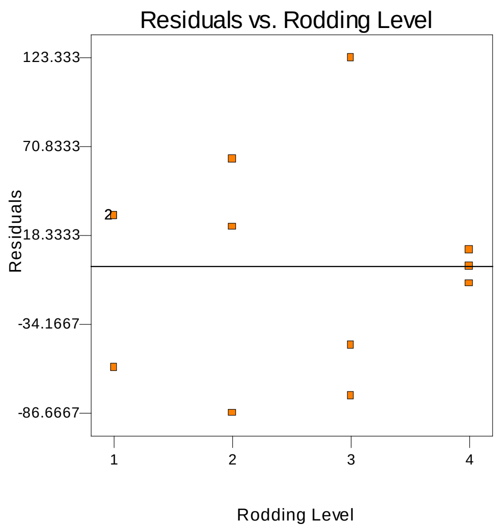
(b) Find the *P*-value for the *F* statistic in part (a). From computer output, P=0.2138.

(c) Analyze the residuals from this experiment. What conclusions can you draw about the underlying model assumptions?

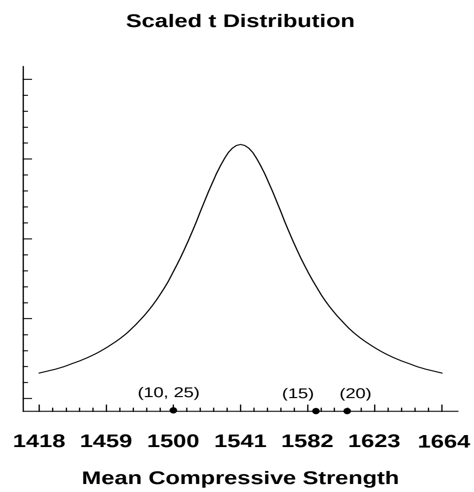
Slight inequality of variance can be observed in the residual plots below; however, not enough to be concerned about the assumptions.







(d) Construct a graphical display to compare the treatment means as describe in Section 3.5.3.



3.21. An article in *Environment International* (Vol. 18, No. 4, 1992) describes an experiment in which the amount of radon released in showers was investigated. Radon enriched water was used in the experiment and six different orifice diameters were tested in shower heads. The data from the experiment are shown in the following table.

Orifice Dia.	Radon Released (%)			
0.37	80	83	83	85
0.51	75	75	79	79
0.71	74	73	76	77
1.02	67	72	74	74
1.40	62	62	67	69
1.99	60	61	64	66

(a) Does the size of the orifice affect the mean percentage of radon released? Use $\alpha = 0.05$.

Yes. There is at least one treatment mean that is different.

Design Expert Output

Response: Radon Released in %
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	1133.38	5	226.68	30.85	< 0.0001	significant
A	*1133.38*	*5*	*226.68*	*30.85*	*< 0.0001*	
Residual	132.25	18	7.35			
Lack of Fit	*0.000*	*0*				
Pure Error	*132.25*	*18*	*7.35*			
Cor Total	1265.63	23				

The Model F-value of 30.85 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-0.37	82.75	1.36
2-0.51	77.00	1.36
3-0.71	75.00	1.36
4-1.02	71.75	1.36
5-1.40	65.00	1.36
6-1.99	62.75	1.36

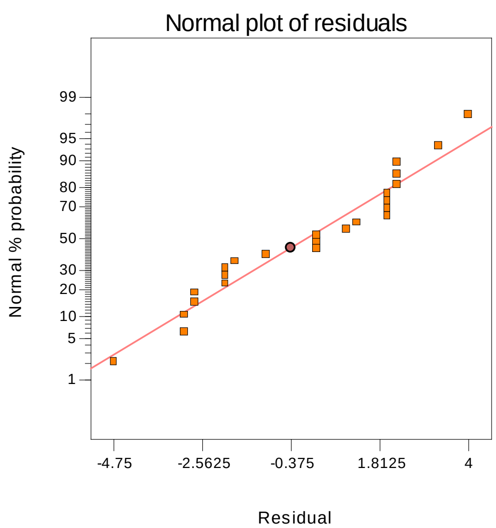
Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	5.75	1	1.92	3.00	0.0077
1 vs 3	7.75	1	1.92	4.04	0.0008
1 vs 4	11.00	1	1.92	5.74	< 0.0001
1 vs 5	17.75	1	1.92	9.26	< 0.0001
1 vs 6	20.00	1	1.92	10.43	< 0.0001
2 vs 3	2.00	1	1.92	1.04	0.3105
2 vs 4	5.25	1	1.92	2.74	0.0135
2 vs 5	12.00	1	1.92	6.26	< 0.0001
2 vs 6	14.25	1	1.92	7.43	< 0.0001
3 vs 4	3.25	1	1.92	1.70	0.1072
3 vs 5	10.00	1	1.92	5.22	< 0.0001
3 vs 6	12.25	1	1.92	6.39	< 0.0001
4 vs 5	6.75	1	1.92	3.52	0.0024
4 vs 6	9.00	1	1.92	4.70	0.0002
5 vs 6	2.25	1	1.92	1.17	0.2557

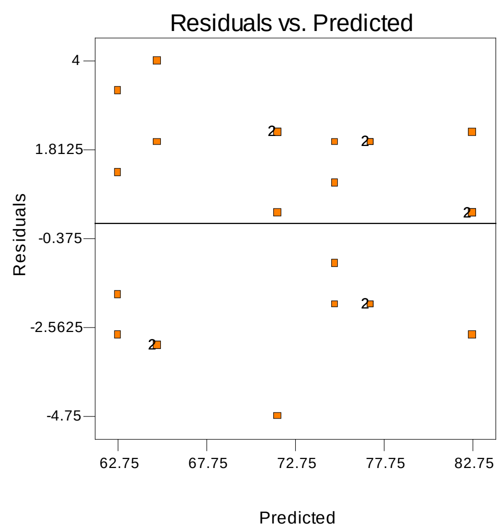
(b) Find the P-value for the F statistic in part (a).

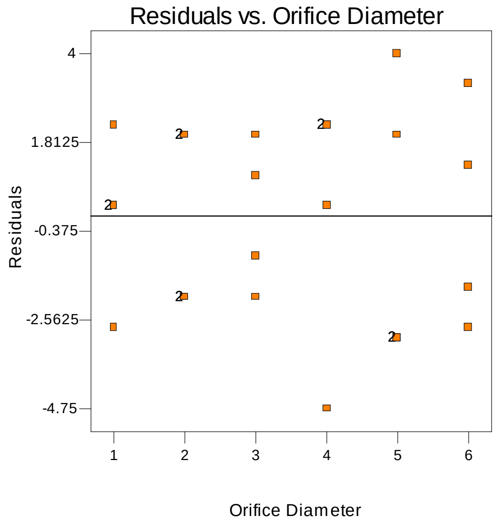
P=3.161 x 10^{-8}

(c) Analyze the residuals from this experiment.

There is nothing unusual about the residuals.





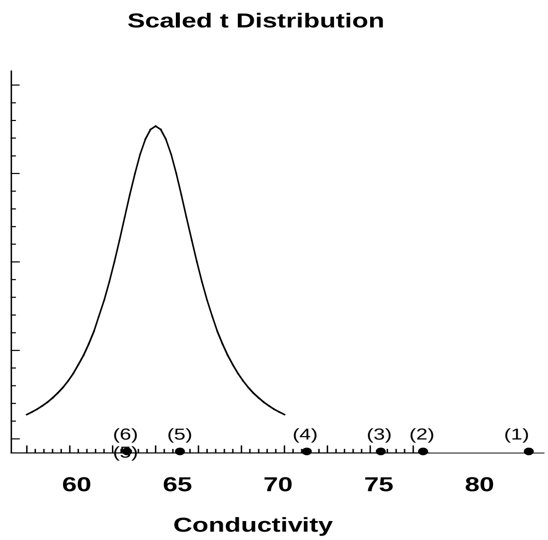


(d) Find a 95 percent confidence interval on the mean percent radon released when the orifice diameter is 1.40.

$$\text{Treatment 5 (Orifice =1.40):} \quad 65 \pm 2.101\sqrt{\frac{7.35}{4}}$$

$$62.152 \le \mu \le 67.848$$

(e) Construct a graphical display to compare the treatment means as describe in Section 3.5.3. What conclusions can you draw?



Treatments 5 and 6 as a group differ from the other means; 2, 3, and 4 as a group differ from the other means, 1 differs from the others.

3.22. The response time in milliseconds was determined for three different types of circuits that could be used in an automatic valve shutoff mechanism. The results are shown in the following table.

Circuit Type	Response Time				
1	9	12	10	8	15
2	20	21	23	17	30
3	6	5	8	16	7

(a) Test the hypothesis that the three circuit types have the same response time. Use $\alpha = 0.01$.

From the computer printout, F=16.08, so there is at least one circuit type that is different.

Design Expert Output

Response: Response Time in ms
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	543.60	2	271.80	16.08	0.0004	significant
A	*543.60*	*2*	*271.80*	*16.08*	*0.0004*	
Residual	202.80	12	16.90			
Lack of Fit	*0.000*	*0*				
Pure Error	*202.80*	*12*	*16.90*			
Cor Total	746.40	14				

The Model F-value of 16.08 implies the model is significant. There is only a 0.04% chance that a "Model F-Value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	10.80	1.84
2-2	22.20	1.84
3-3	8.40	1.84

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-11.40	1	2.60	-4.38	0.0009
1 vs 3	2.40	1	2.60	0.92	0.3742
2 vs 3	13.80	1	2.60	5.31	0.0002

(b) Use Tukey's test to compare pairs of treatment means. Use $\alpha = 0.01$.

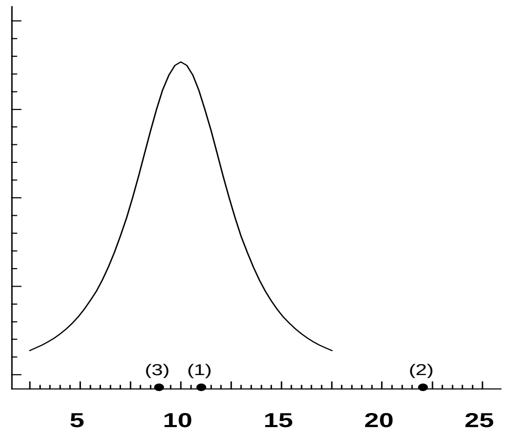
$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{n}} = \sqrt{\frac{16.90}{5}} = 1.8385$$
$$q_{0.01,(3,12)} = 5.04$$
$$t_0 = 1.8385(5.04) = 9.266$$
$$\text{1 vs. 2: } |10.8\text{-}22.2| = 11.4 > 9.266$$
$$\text{1 vs. 3: } |10.8\text{-}8.4| = 2.4 < 9.266$$
$$\text{2 vs. 3: } |22.2\text{-}8.4| = 13.8 > 9.266$$
1 and 2 are different. 2 and 3 are different.

Notice that the results indicate that the mean of treatment 2 differs from the means of both treatments 1 and 3, and that the means for treatments 1 and 3 are the same. Notice also that the Fisher LSD procedure (see the computer output) gives the same results.

(c) Use the graphical procedure in Section 3.5.3 to compare the treatment means. What conclusions can you draw? How do they compare with the conclusions from part (a).

The scaled-*t* plot agrees with part (b). In this case, the large difference between the mean of treatment 2 and the other two treatments is very obvious.

Scaled t Distribution



(d) Construct a set of orthogonal contrasts, assuming that at the outset of the experiment you suspected the response time of circuit type 2 to be different from the other two.

$$H_0 = \mu_1 - 2\mu_2 + \mu_3 = 0$$
$$H_1 = \mu_1 - 2\mu_2 + \mu_3 \neq 0$$
$$C_1 = y_{1.} - 2y_{2.} + y_{3.}$$
$$C_1 = 54 - 2(111) + 42 = -126$$
$$SS_{C1} = \frac{(-126)^2}{5(6)} = 529.2$$
$$F_{C1} = \frac{529.2}{16.9} = 31.31$$

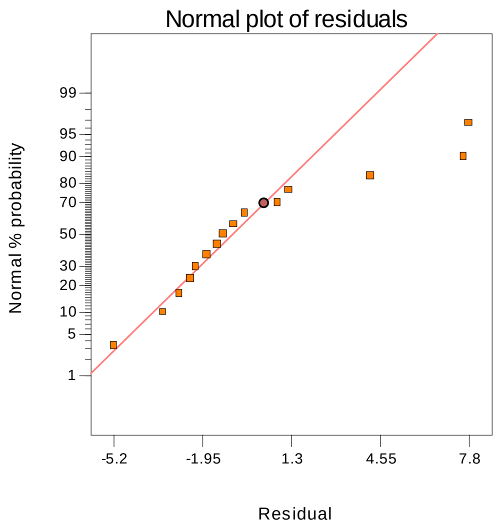
Type 2 differs from the average of type 1 and type 3.

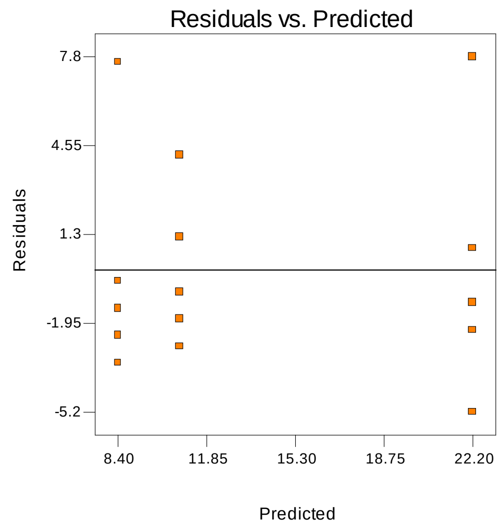
(e) If you were a design engineer and you wished to minimize the response time, which circuit type would you select?

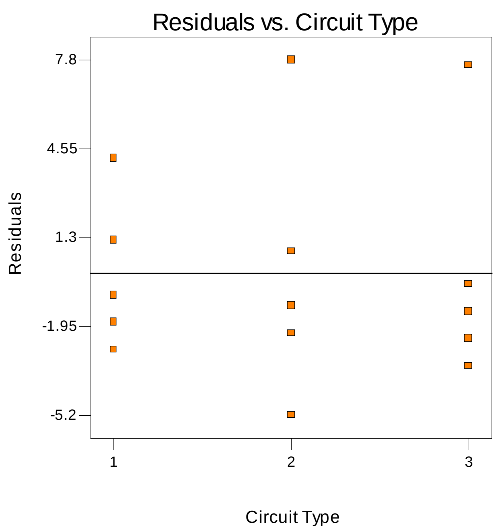
Either type 1 or type 3 as they are not different from each other and have the lowest response time.

(f) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied?

The normal probability plot has some points that do not lie along the line in the upper region. This may indicate potential outliers in the data.







3.23. The effective life of insulating fluids at an accelerated load of 35 kV is being studied. Test data have been obtained for four types of fluids. The results from a completely randomized experiment were as follows:

Fluid Type	Life (in h) at 35 kV Load					
1	17.6	18.9	16.3	17.4	20.1	21.6
2	16.9	15.3	18.6	17.1	19.5	20.3
3	21.4	23.6	19.4	18.5	20.5	22.3
4	19.3	21.1	16.9	17.5	18.3	19.8

(a) Is there any indication that the fluids differ? Use $\alpha = 0.05$.

At $\alpha = 0.05$ there is no difference, but since the *P*-value is just slightly above 0.05, there is probably a difference in means.

Design Expert Output

Response: Life in in h
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	30.17	3	10.06	3.05	0.0525	not significant
A	*30.16*	*3*	*10.05*	*3.05*	*0.0525*	
Residual	65.99	20	3.30			
Lack of Fit	*0.000*	*0*				
Pure Error	*65.99*	*20*	*3.30*			
Cor Total	96.16	23				

The Model F-value of 3.05 implies there is a 5.25% chance that a "Model F-Value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

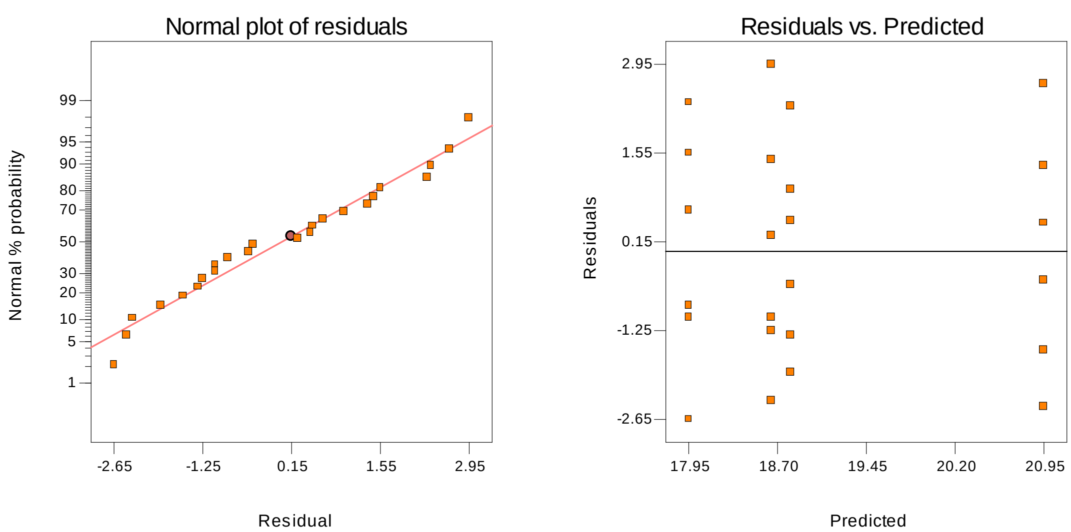
	Estimated Mean	**Standard Error**
1-1	18.65	0.74
2-2	17.95	0.74
3-3	20.95	0.74
4-4	18.82	0.74

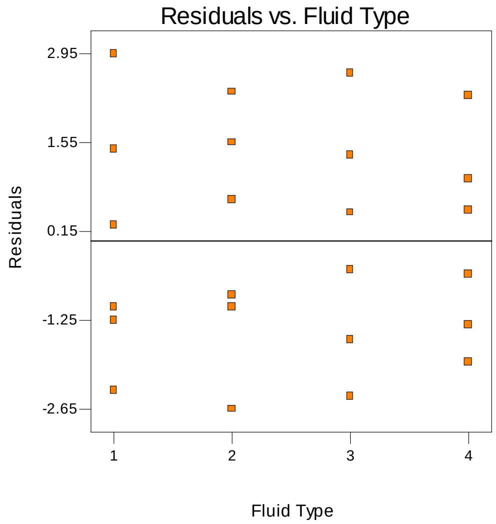
Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H_0 Coeff=0**	**Prob > |t|**
1 vs 2	0.70	1	1.05	0.67	0.5121
1 vs 3	-2.30	1	1.05	-2.19	0.0403
1 vs 4	-0.17	1	1.05	-0.16	0.8753
2 vs 3	-3.00	1	1.05	-2.86	0.0097
2 vs 4	-0.87	1	1.05	-0.83	0.4183
3 vs 4	2.13	1	1.05	2.03	0.0554

(b) Which fluid would you select, given that the objective is long life?

Treatment 3. The Fisher LSD procedure in the computer output indicates that the fluid 3 is different from the others, and it's average life also exceeds the average lives of the other three fluids.

(c) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied? There is nothing unusual in the residual plots.





3.24. Four different designs for a digital computer circuit are being studied in order to compare the amount of noise present. The following data have been obtained:

Circuit Design	Noise Observed				
1	19	20	19	30	8
2	80	61	73	56	80
3	47	26	25	35	50
4	95	46	83	78	97

(a) Is the amount of noise present the same for all four designs? Use $\alpha = 0.05$.

No, at least one treatment mean is different.

Design Expert Output

Response: Noise
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	12042.00	3	4014.00	21.78	< 0.0001	significant
A	*12042.00*	*3*	*4014.00*	*21.78*	*< 0.0001*	
Residual	2948.80	16	184.30			
Lack of Fit	*0.000*	*0*				
Pure Error	*2948.80*	*16*	*184.30*			
Cor Total	14990.80	19				

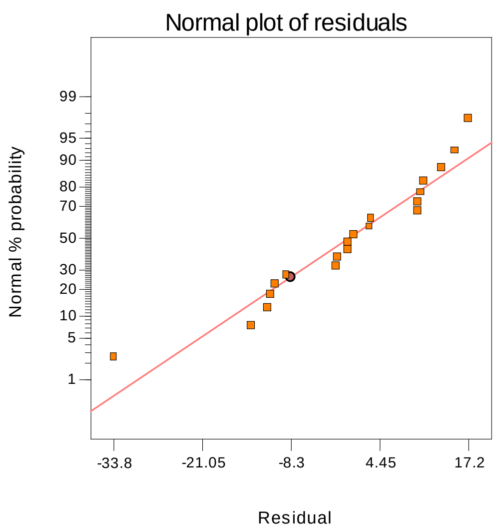
The Model F-value of 21.78 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

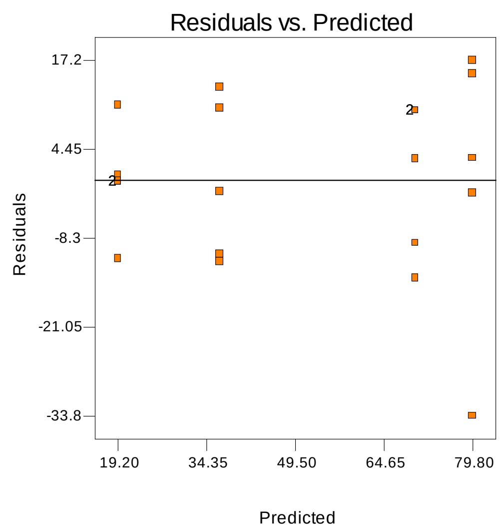
Treatment Means (Adjusted, If Necessary)

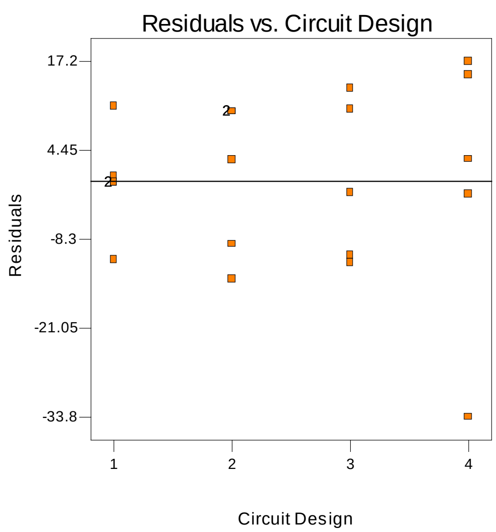
	Estimated Mean	**Standard Error**
1-1	19.20	6.07
2-2	70.00	6.07
3-3	36.60	6.07
4-4	79.80	6.07

Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	-50.80	1	8.59	-5.92	< 0.0001
1 vs 3	-17.40	1	8.59	-2.03	0.0597
1 vs 4	-60.60	1	8.59	-7.06	< 0.0001
2 vs 3	33.40	1	8.59	3.89	0.0013
2 vs 4	-9.80	1	8.59	-1.14	0.2705
3 vs 4	-43.20	1	8.59	-5.03	0.0001

(b) Analyze the residuals from this experiment. Are the basic analysis of variance assumptions satisfied?
There is nothing too unusual about the residual plots, although there is a mild outlier present.







(c) Which circuit design would you select for use? Low noise is best.

From the Design Expert Output, the Fisher LSD procedure comparing the difference in means identifies Type 1 as having lower noise than Types 2 and 4. Although the LSD procedure comparing Types 1 and 3 has a *P*-value greater than 0.05, it is less than 0.10. Unless there are other reasons for choosing Type 3, Type 1 would be selected.

3.25. Four chemists are asked to determine the percentage of methyl alcohol in a certain chemical compound. Each chemist makes three determinations, and the results are the following:

Chemist	Percentage of Methyl Alcohol		
1	84.99	84.04	84.38
2	85.15	85.13	84.88
3	84.72	84.48	85.16
4	84.20	84.10	84.55

(a) Do chemists differ significantly? Use $\alpha = 0.05$.

There is no significant difference at the 5% level, but chemists differ significantly at the 10% level.

Design Expert Output

Response: Methyl Alcohol in %
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	1.04	3	0.35	3.25	0.0813	not significant
A	*1.04*	*3*	*0.35*	*3.25*	*0.0813*	
Residual	0.86	8	0.11			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.86*	*8*	*0.11*			
Cor Total	1.90	11				

The Model F-value of 3.25 implies there is a 8.13% chance that a "Model F-Value" this large could occur due to noise.

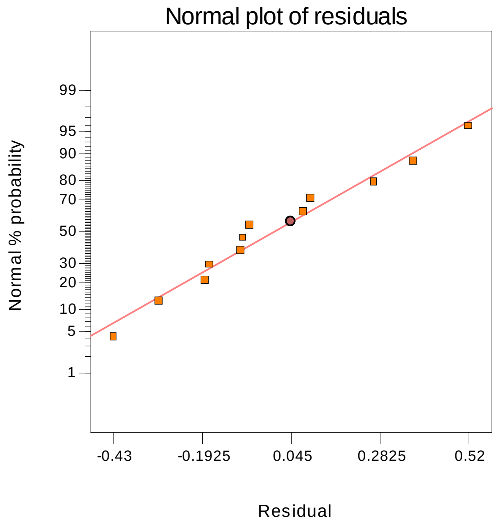
Treatment Means (Adjusted, If Necessary)

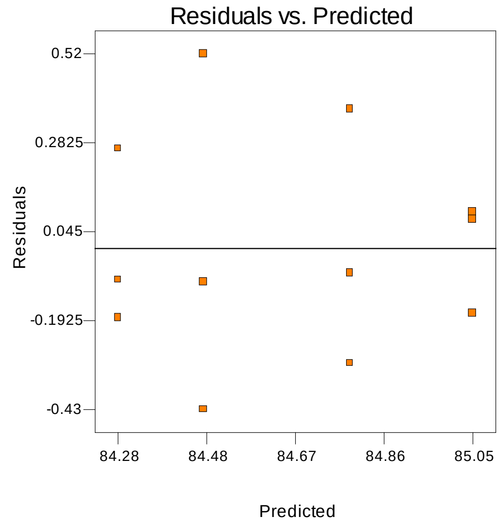
	Estimated Mean	Standard Error
1-1	84.47	0.19
2-2	85.05	0.19
3-3	84.79	0.19
4-4	84.28	0.19

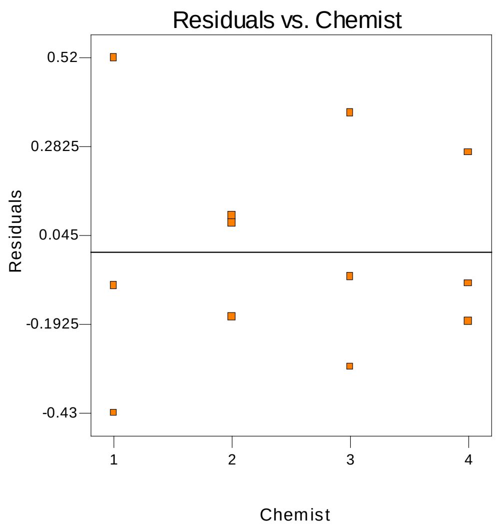
Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	-0.58	1	0.27	-2.18	0.0607
1 vs 3	-0.32	1	0.27	-1.18	0.2703
1 vs 4	0.19	1	0.27	0.70	0.5049
2 vs 3	0.27	1	0.27	1.00	0.3479
2 vs 4	0.77	1	0.27	2.88	0.0205
3 vs 4	0.50	1	0.27	1.88	0.0966

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residual plots.







(c) If chemist 2 is a new employee, construct a meaningful set of orthogonal contrasts that might have been useful at the start of the experiment.

Chemists	Total	C1	C2	C3
1	253.41	1	-2	0
2	255.16	-3	0	0
3	254.36	1	1	-1
4	252.85	1	1	1
	Contrast Totals:	-4.86	0.39	-1.51

$$SS_{C1} = \frac{(-4.86)^2}{3(12)} = 0.656 \quad F_{C1} = \frac{0.656}{0.10727} = 6.115*$$
$$SS_{C2} = \frac{(0.39)^2}{3(6)} = 0.008 \quad F_{C2} = \frac{0.008}{0.10727} = 0.075$$
$$SS_{C3} = \frac{(-1.51)^2}{3(2)} = 0.380 \quad F_{C3} = \frac{0.380}{0.10727} = 3.54$$

Only contrast 1 is significant at 5%.

3.26. Three brands of batteries are under study. It is s suspected that the lives (in weeks) of the three brands are different. Five randomly selected batteries of each brand are tested with the following results:

Weeks of Life		
Brand 1	Brand 2	Brand 3
100	76	108
96	80	100
92	75	96
96	84	98
92	82	100

(a) Are the lives of these brands of batteries different?

Yes, at least one of the brands is different.

Design Expert Output

Response: Life in Weeks
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	1196.13	2	598.07	38.34	< 0.0001	significant
A	*1196.13*	*2*	*598.07*	*38.34*	*< 0.0001*	
Residual	187.20	12	15.60			
Lack of Fit	*0.000*	*0*				
Pure Error	*187.20*	*12*	*15.60*			
Cor Total	1383.33	14				

The Model F-value of 38.34 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

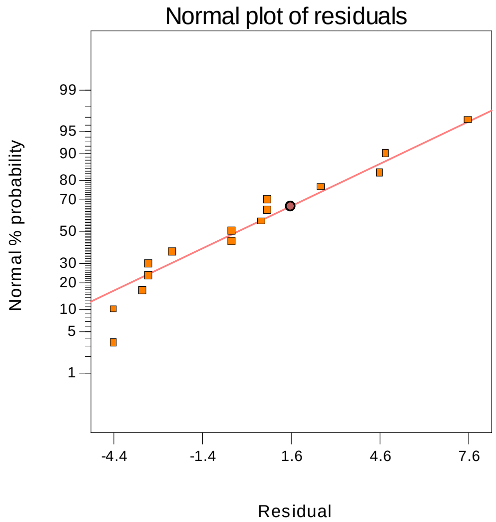
Treatment Means (Adjusted, If Necessary)

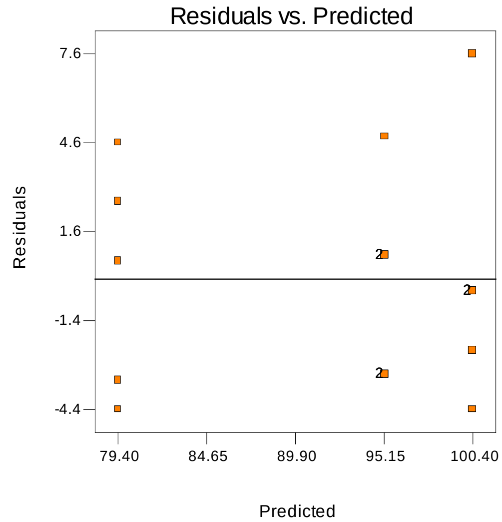
	Estimated Mean	**Standard Error**
1-1	95.20	1.77
2-2	79.40	1.77
3-3	100.40	1.77

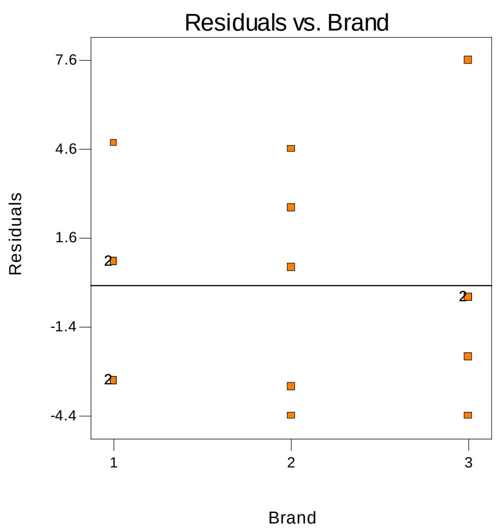
Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	15.80	1	2.50	6.33	< 0.0001
1 vs 3	-5.20	1	2.50	-2.08	0.0594
2 vs 3	-21.00	1	2.50	-8.41	< 0.0001

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residuals.







(c) Construct a 95 percent interval estimate on the mean life of battery brand 2. Construct a 99 percent interval estimate on the mean difference between the lives of battery brands 2 and 3.

$$\bar{y}_{i.} \pm t_{\alpha/2, N-a}\sqrt{\frac{MS_E}{n}}$$

$$\text{Brand 2:}\quad 79.4 \pm 2.179\sqrt{\frac{15.60}{5}}$$

$$79.40 \pm 3.849$$

$$75.551 \le \mu_2 \le 83.249$$

$$\text{Brand 2 - Brand 3:}\quad \bar{y}_{i.} - \bar{y}_{j.} \pm t_{\alpha/2, N-a}\sqrt{\frac{2MS_E}{n}}$$

$$79.4 - 100.4 \pm 3.055\sqrt{\frac{2(15.60)}{5}}$$

$$-28.631 \le \mu_2 - \mu_3 \le -13.369$$

(d) Which brand would you select for use? If the manufacturer will replace without charge any battery that fails in less than 85 weeks, what percentage would the company expect to replace?

Chose brand 3 for longest life. Mean life of this brand in 100.4 weeks, and the variance of life is estimated by 15.60 (*MSE*). Assuming normality, then the probability of failure before 85 weeks is:

$$\Phi\left(\frac{85-100.4}{\sqrt{15.60}}\right)=\Phi(-3.90)=0.00005$$

That is, about 5 out of 100,000 batteries will fail before 85 week.

3.27. Four catalysts that may affect the concentration of one component in a three component liquid mixture are being investigated. The following concentrations are obtained from a completely randomized experiment:

Catalyst			
1	2	3	4
58.2	56.3	50.1	52.9
57.2	54.5	54.2	49.9
58.4	57.0	55.4	50.0
55.8	55.3		51.7
54.9			

(a) Do the four catalysts have the same effect on concentration?

No, their means are different.

Design Expert Output

Response: Concentration

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	85.68	3	28.56	9.92	0.0014	significant
A	*85.68*	*3*	*28.56*	*9.92*	*0.0014*	
Residual	34.56	12	2.88			
Lack of Fit	*0.000*	*0*				
Pure Error	*34.56*	*12*	*2.88*			
Cor Total	120.24	15				

The Model F-value of 9.92 implies the model is significant. There is only a 0.14% chance that a "Model F-Value" this large could occur due to noise.

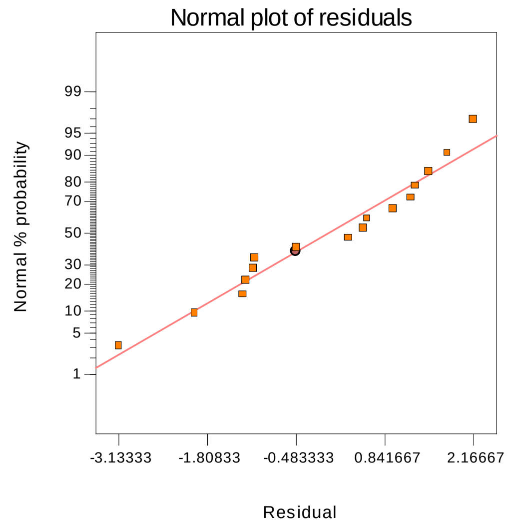
Treatment Means (Adjusted, If Necessary)

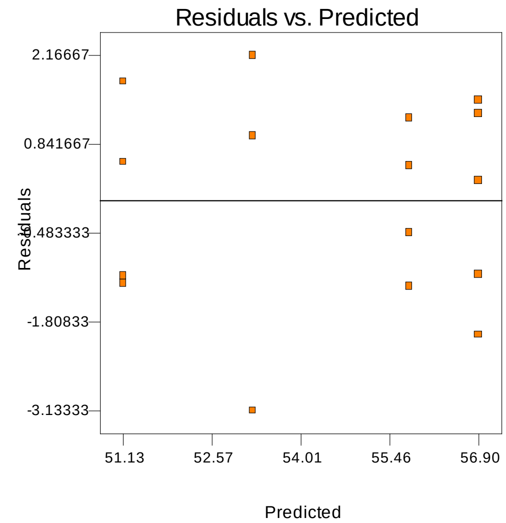
	Estimated Mean	**Standard Error**
1-1	56.90	0.76
2-2	55.77	0.85
3-3	53.23	0.98
4-4	51.13	0.85

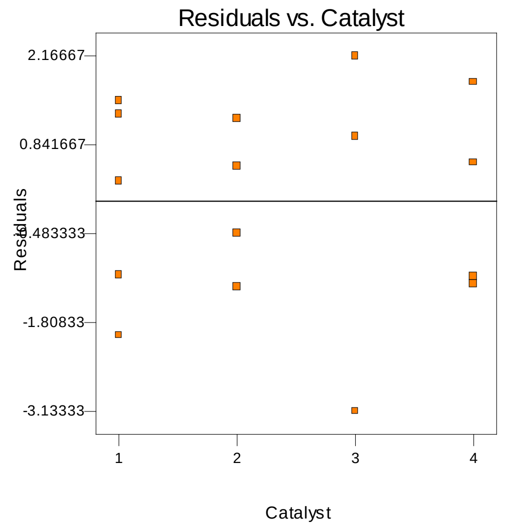
Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	1.13	1	1.14	0.99	0.3426
1 vs 3	3.67	1	1.24	2.96	0.0120
1 vs 4	5.77	1	1.14	5.07	0.0003
2 vs 3	2.54	1	1.30	1.96	0.0735
2 vs 4	4.65	1	1.20	3.87	0.0022
3 vs 4	2.11	1	1.30	1.63	0.1298

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residual plots.







(c) Construct a 99 percent confidence interval estimate of the mean response for catalyst 1.

$$\bar{y}_{i.} \pm t_{\alpha/2, N-a}\sqrt{\frac{MS_E}{n}}$$

$$\text{Catalyst 1:} \quad 56.9 \pm 3.055\sqrt{\frac{2.88}{5}}$$

$$56.9 \pm 2.3186$$

$$54.5814 \le \mu_1 \le 59.2186$$

3.28. An experiment was performed to investigate the effectiveness of five insulating materials. Four samples of each material were tested at an elevated voltage level to accelerate the time to failure. The failure times (in minutes) is shown below:

Material	Failure Time (minutes)			
1	110	157	194	178
2	1	2	4	18
3	880	1256	5276	4355
4	495	7040	5307	10050
5	7	5	29	2

(a) Do all five materials have the same effect on mean failure time?

No, at least one material is different.

Design Expert Output

Response: Failure Timein Minutes
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	1.032E+008	4	2.580E+007	6.19	0.0038	significant
A	*1.032E+008*	*4*	*2.580E+007*	*6.19*	*0.0038*	
Residual	6.251E+007	15	4.167E+006			
Lack of Fit	*0.000*	*0*				
Pure Error	*6.251E+007*	*15*	*4.167E+006*			
Cor Total	1.657E+008	19				

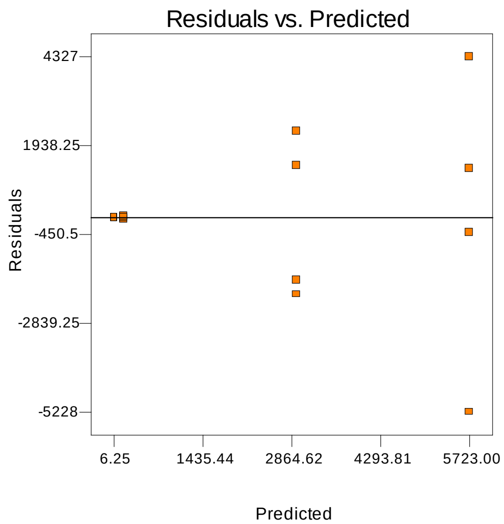
The Model F-value of 6.19 implies the model is significant. There is only a 0.38% chance that a "Model F-Value" this large could occur due to noise.

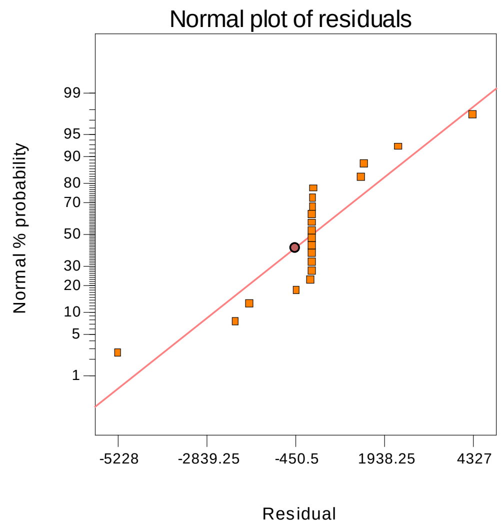
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	159.75	1020.67
2-2	6.25	1020.67
3-3	2941.75	1020.67
4-4	5723.00	1020.67
5-5	10.75	1020.67

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	153.50	1	1443.44	0.11	0.9167
1 vs 3	-2782.00	1	1443.44	-1.93	0.0731
1 vs 4	-5563.25	1	1443.44	-3.85	0.0016
1 vs 5	149.00	1	1443.44	0.10	0.9192
2 vs 3	-2935.50	1	1443.44	-2.03	0.0601
2 vs 4	-5716.75	1	1443.44	-3.96	0.0013
2 vs 5	-4.50	1	1443.44	-3.118E-003	0.9976
3 vs 4	-2781.25	1	1443.44	-1.93	0.0732
3 vs 5	2931.00	1	1443.44	2.03	0.0604
4 vs 5	5712.25	1	1443.44	3.96	0.0013

(b) Plot the residuals versus the predicted response. Construct a normal probability plot of the residuals. What information do these plots convey?





The plot of residuals versus predicted has a strong outward-opening funnel shape, which indicates the variance of the original observations is not constant. The normal probability plot also indicates that the normality assumption is not valid. A data transformation is recommended.

(c) Based on your answer to part (b) conduct another analysis of the failure time data and draw appropriate conclusions.

A natural log transformation was applied to the failure time data. The analysis in the log scale identifies that there exists at least one difference in treatment means.

Design Expert Output

Response: Failure Timein Minutes 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	165.06	4	41.26	37.66	< 0.0001	significant
A	*165.06*	*4*	*41.26*	*37.66*	*< 0.0001*	
Residual	16.44	15	1.10			
Lack of Fit	*0.000*	*0*				
Pure Error	*16.44*	*15*	*1.10*			
Cor Total	181.49	19				

The Model F-value of 37.66 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

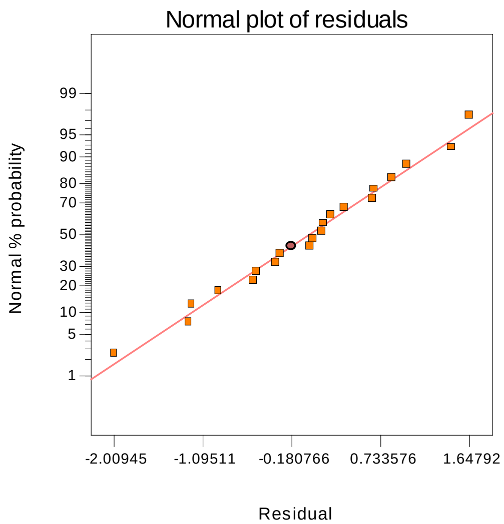
Treatment Means (Adjusted, If Necessary)

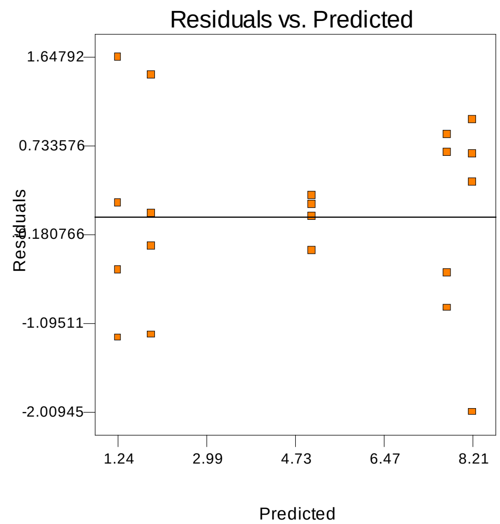
	Estimated Mean	**Standard Error**
1-1	5.05	0.52
2-2	1.24	0.52
3-3	7.72	0.52
4-4	8.21	0.52
5-5	1.90	0.52

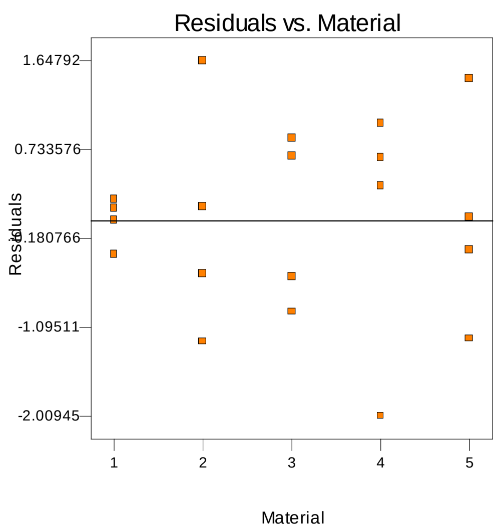
Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	3.81	1	0.74	5.15	0.0001
1 vs 3	-2.66	1	0.74	-3.60	0.0026
1 vs 4	-3.16	1	0.74	-4.27	0.0007
1 vs 5	3.15	1	0.74	4.25	0.0007
2 vs 3	-6.47	1	0.74	-8.75	< 0.0001
2 vs 4	-6.97	1	0.74	-9.42	< 0.0001

2 vs 5	-0.66	1	0.74	-0.89	0.3856
3 vs 4	-0.50	1	0.74	-0.67	0.5116
3 vs 5	5.81	1	0.74	7.85	< 0.0001
4 vs 5	6.31	1	0.74	8.52	< 0.0001

There is nothing unusual about the residual plots when the natural log transformation is applied.







3.29. A semiconductor manufacturer has developed three different methods for reducing particle counts on wafers. All three methods are tested on five wafers and the after-treatment particle counts obtained. The data are shown below:

Method	Count				
1	31	10	21	4	1
2	62	40	24	30	35
3	58	27	120	97	68

(a) Do all methods have the same effect on mean particle count?

No, at least one method has a different effect on mean particle count.

Design Expert Output

Response: **Count**
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	8963.73	2	4481.87	7.91	0.0064	significant
A	*8963.73*	*2*	*4481.87*	*7.91*	*0.0064*	
Residual	6796.00	12	566.33			
Lack of Fit	*0.000*	*0*				
Pure Error	*6796.00*	*12*	*566.33*			
Cor Total	15759.73	14				

The Model F-value of 7.91 implies the model is significant. There is only a 0.64% chance that a "Model F-Value" this large could occur due to noise.

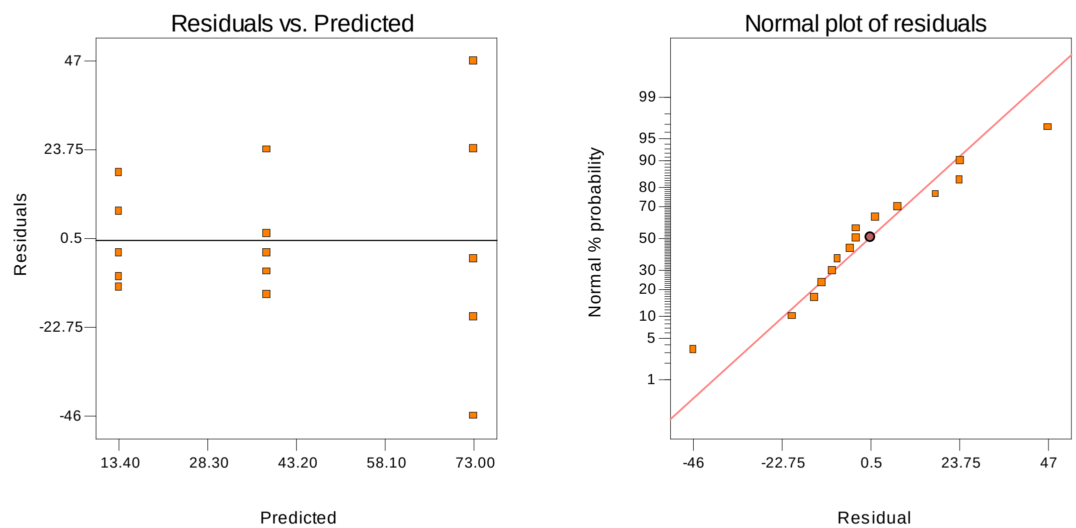
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	13.40	10.64
2-2	38.20	10.64
3-3	73.00	10.64

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-24.80	1	15.05	-1.65	0.1253
1 vs 3	-59.60	1	15.05	-3.96	0.0019
2 vs 3	-34.80	1	15.05	-2.31	0.0393

(b) Plot the residuals versus the predicted response. Construct a normal probability plot of the residuals. Are there potential concerns about the validity of the assumptions?

The plot of residuals versus predicted appears to be funnel shaped. This indicates the variance of the original observations is not constant. The residuals plotted in the normal probability plot do not fall along a straight line, which suggests that the normality assumption is not valid. A data transformation is recommended.



(c) Based on your answer to part (b) conduct another analysis of the particle count data and draw appropriate conclusions.

For count data, a square root transformation is often very effective in resolving problems with inequality of variance. The analysis of variance for the transformed response is shown below. The difference between methods is much more apparent after applying the square root transformation.

Design Expert Output

Response: Count Transform: Square root 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	63.90	2	31.95	9.84	0.0030	significant
A	*63.90*	*2*	*31.95*	*9.84*	*0.0030*	
Residual	38.96	12	3.25			
Lack of Fit	*0.000*	*0*				
Pure Error	*38.96*	*12*	*3.25*			
Cor Total	102.86	14				

The Model F-value of 9.84 implies the model is significant. There is only a 0.30% chance that a "Model F-Value" this large could occur due to noise.

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	3.26	0.81
2-2	6.10	0.81
3-3	8.31	0.81

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-2.84	1	1.14	-2.49	0.0285
1 vs 3	-5.04	1	1.14	-4.42	0.0008
2 vs 3	-2.21	1	1.14	-1.94	0.0767

3.30 A manufacturer suspects that the batches of raw material furnished by his supplier differ significantly in calcium content. There are a large number of batches currently in the warehouse. Five of these are randomly selected for study. A chemist makes five determinations on each batch as obtains the following data:

Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
23.46	23.59	23.51	23.28	23.29
23.48	23.46	23.64	23.40	23.46
23.56	23.42	23.46	23.37	23.37
23.39	23.49	23.52	23.46	23.32
23.40	23.50	23.49	23.39	23.38

(a) Is there significant variation in the calcium content from batch to batch? Use α=0.05. The computer output below shows that for the random effects model there is batch to batch variation.

Based on the ANOVA in the JMP output below, the batches differ significantly.

JMP Output

Summary of Fit

RSquare	0.525399
RSquare Adj	0.430479
Root Mean Square Error	0.066182
Mean of Response	23.4436
Observations (or Sum Wgts)	25

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	0.09697600	0.024244	5.5352
Error	20	0.08760000	0.004380	**Prob > F**
C. Total	24	0.18457600		0.0036*

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Batch	4	4	0.09697600	5.5352	0.0036*

(b) Estimate the components of variance.

$$\hat{\sigma}^2 = MS_E = 0.004380$$
$$\hat{\sigma}_\tau^2 = \frac{MS_{Treatments} - MS_E}{n} = \frac{0.024244 - 0.004380}{5} = 0.003973$$

This is verified in the JMP REML analysis shown below.

JMP Output

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	23.4436	0.031141	4	752.82	<.0001*

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Batch	0.907032	0.0039728	0.0034398	-0.002769	0.0107147	47.562
Residual		0.00438	0.0013851	0.0025637	0.0091338	52.438
Total		0.0083528				100.000

Covariance Matrix of Variance Component Estimates

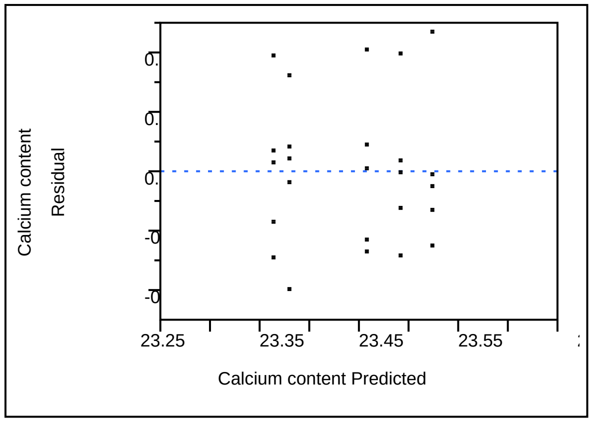
Random Effect	Batch	Residual
Batch	1.1832e-5	-3.837e-7
Residual	-3.837e-7	1.9184e-6

(c) Find a 95 percent confidence interval for $\sigma_\tau^2 \big/ \left(\sigma_\tau^2 + \sigma^2\right)$

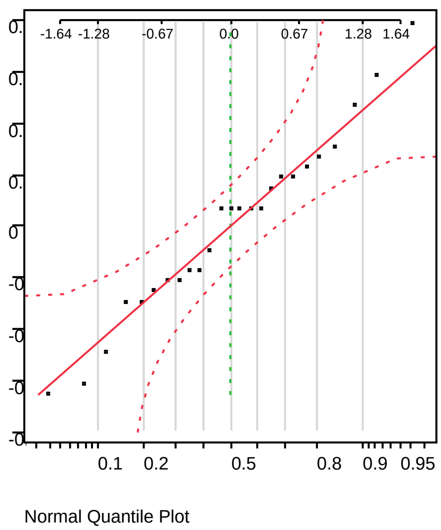
$$L = \frac{1}{n}\left(\frac{MS_{Treatments}}{MS_E}\frac{1}{F_{\alpha/2,a-1,N-a}} - 1\right) = \frac{1}{5}\left(\frac{0.024244}{0.004380}\frac{1}{3.51} - 1\right) = 0.1154$$
$$U = \frac{1}{n}\left(\frac{MS_{Treatments}}{MS_E}\frac{1}{F_{1-\alpha/2,a-1,N-a}} - 1\right) = \frac{1}{5}\left(\frac{0.024244}{0.004380}\frac{1}{0.1168} - 1\right) = 9.2780$$
$$\frac{L}{1+L} \le \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} \le \frac{U}{1+U}$$
$$\frac{0.1154}{1+0.1154} \le \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} \le \frac{9.2780}{1+9.2780}$$
$$0.1035 \le \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} \le 0.9027$$

(d) Analyze the residuals from this experiment. Are the analysis of variance assumptions are satisfied?

The plot of residuals vs. predicted show no concerns.



The residuals used in the plot below are based on the REML analysis and shows no concerns. Note, normality is not a concern for this analysis.



3.31. Several ovens in a metal working shop are used to heat metal specimens. All ovens are supposed to operate at the same temperature, although it is suspected that this may not be true. Three ovens selected at random, and their temperatures on successive heats are noted. The data collected are as follows:

Oven	Temperature					
1	491.50	498.30	498.10	493.50	493.60	
2	488.50	484.65	479.90	477.35		
3	480.10	484.80	488.25	473.00	471.85	478.65

(a) Is there significant variation in temperature between ovens? Use α=0.05.

The computer output below shows that there is oven to oven variation.

Minitab Output

```
ANOVA: Temp versus Oven

Factor  Type    Levels  Values
Oven    random       3  1, 2, 3


Analysis of Variance for Temp

Source      DF        SS        MS        F        P
Oven         2    705.10    352.55    13.33    0.001
Error       12    317.31     26.44
Total       14   1022.41


S = 5.14224   R-Sq = 68.96%   R-Sq(adj) = 63.79%
```

(b) Estimate the components of variation for this model.

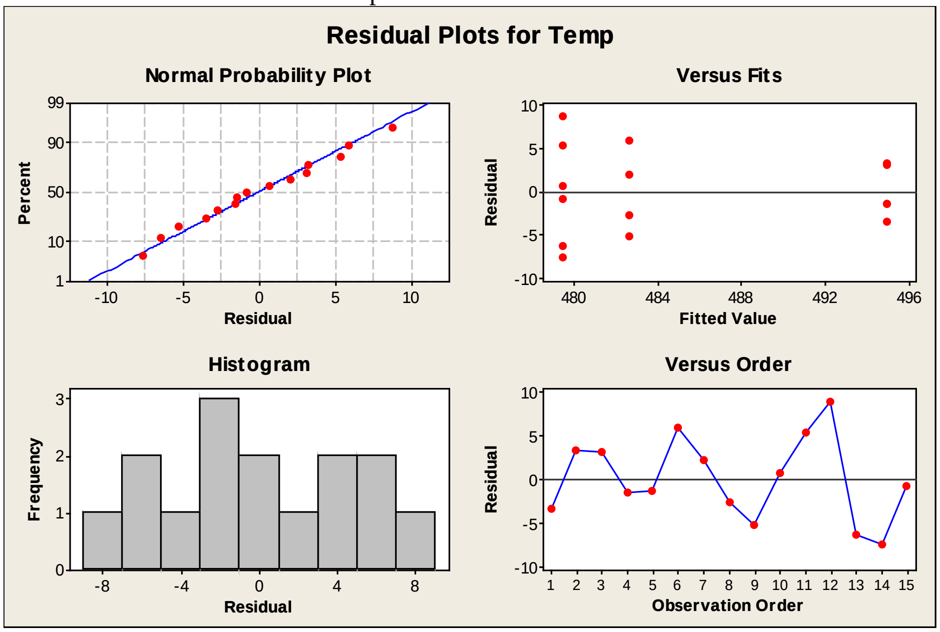
$$n_0 = \frac{1}{a-1}\left[\sum_{i=1}^{a} n_i - \frac{\sum_{i=1}^{a} n_i^2}{\sum_{i=1}^{a} n_i}\right] = \frac{1}{3-1}\left[15 - \frac{77}{15}\right] = 4.9333$$

$$\hat{\sigma}^2 = MS_E = 26.44$$

$$\hat{\sigma}_\tau^2 = \frac{MS_{Treatments} - MS_E}{n_0} = \frac{352.55 - 26.44}{4.9333} = 66.10$$

(c) Analyze the residuals from this experiment and draw conclusions about model adequacy.

There are no concerns with the residual plots below.



3.32. An article in the *Journal of the Electrochemical Society* (Vol. 139, No. 2, 1992, pp. 524-532) describes an experiment to investigate low-pressure vapor deposition of polysilicon. The experiment was carried out in a large capacity reactor at Sematech in Austin, Texas. The reactor has several wafer positions, and four of these positions are selected at random. The response variable is film thickness uniformity. Three replicates of the experiment were run, and the data are as follows:

Wafer Positions	Uniformity		
1	2.76	5.67	4.49
2	1.43	1.70	2.19
3	2.34	1.97	1.47
4	0.94	1.36	1.65

(a) Is there a difference in the wafer positions? Use Use α=0.05.

The JMP output below identifies a difference in the wafer positions.

JMP Output

Summary of Fit

RSquare	0.756617
RSquare Adj	0.665349
Root Mean Square Error	0.807579
Mean of Response	2.330833
Observations (or Sum Wgts)	12

Analysis of Variance

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Ratio**
Model	3	16.219825	5.40661	8.2900
Error	8	5.217467	0.65218	Prob > F
C. Total	11	21.437292		0.0077*

Effect Tests

Source	**Nparm**	**DF**	**Sum of Squares**	**F Ratio**	**Prob > F**
Wafer Position	3	3	16.219825	8.2900	0.0077*

(b) Estimate the variability due to wafer position.

The JMP REML output below identifies the variance component for the wafer position as 1.5848.

JMP Output

Parameter Estimates

Term	**Estimate**	**Std Error**	**DFDen**	**t Ratio**	**Prob>|t|**
Intercept	2.3308333	0.671231	3	3.47	0.0403*

REML Variance Component Estimates

Random Effect	**Var Ratio**	**Var Component**	**Std Error**	**95% Lower**	**95% Upper**	**Pct of Total**
Wafer Position	2.4300043	1.5848083	1.4755016	-1.307122	4.4767383	70.846
Residual		0.6521833	0.3260917	0.2975536	2.393629	29.154
Total		2.2369917				100.000

Covariance Matrix of Variance Component Estimates

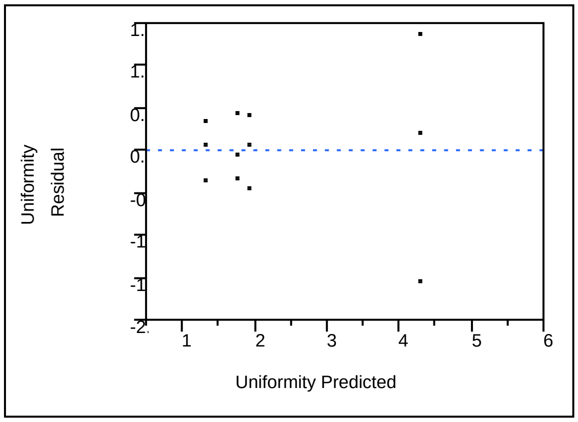
Random Effect	**Wafer Position**	**Residual**
Wafer Position	2.177105	-0.035445
Residual	-0.035445	0.1063358

(c) Estimate the random error component.

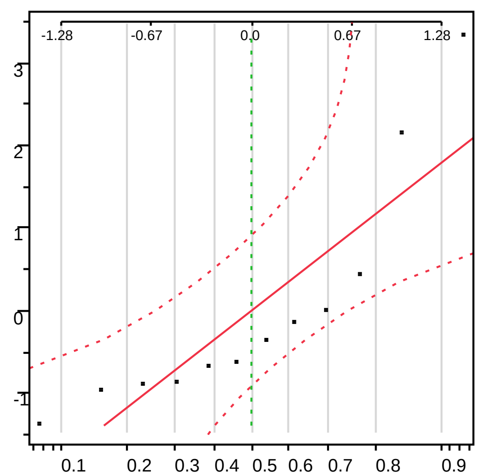
The JMP REML output above identifies the random error variance component as 0.6522..

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

The plot of residuals vs. predicted shows some uniformity concerns.



The residuals used in the plot below are based on the REML analysis. The normal plot shows some concerns with the normality assumption; however, the normality is not important for this analysis.



Normal Quantile Plot

Uniformity data often requires a transformation, such as a log transformation, and should be considered for this experiment.

3.33. Consider the vapor-deposition experiment described in Problem 3.32.

(a) Estimate the total variability in the uniformity response.

The JMP REML output shown in part (b) of Problem 3.32 identifies the total variability as 2.2370.

(b) How much of the total variability in the uniformity response is due to the difference between positions in the reactor?

From the JMP REML output shown in part (b) of Problem 3.32, the differences between positions represents 70.846% of the total variability.

(c) To what level could the variability in the uniformity response be reduced if position-to-position variability in the reactor could be eliminated? Do you believe this is a significant reduction?

The variability could be reduced to 29.154% of the current total variability. Based on the 95% confidence intervals calculated below, this is not significant. An increase in sample size might reverse this decision.

$$L = \frac{1}{n}\left(\frac{MS_{Treatments}}{MS_E}\frac{1}{F_{\alpha/2,a-1,N-a}} - 1\right) = \frac{1}{4}\left(\frac{5.40661}{0.65218}\frac{1}{6.059467} - 1\right) = 0.073623$$

$$U = \frac{1}{n}\left(\frac{MS_{Treatments}}{MS_E}\frac{1}{F_{1-\alpha/2,a-1,N-a}} - 1\right) = \frac{1}{4}\left(\frac{5.40661}{0.65218}\frac{1}{0.025398} - 1\right) = 65.08093$$

$$\frac{L}{1+L} \le \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} \le \frac{U}{1+U}$$

$$\frac{0.073623}{1+0.073623} \le \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} \le \frac{65.08093}{1+65.08093}$$

$$0.068575 \le \frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2 + \hat{\sigma}^2} \le 0.984867$$

3.34. An article in the *Journal of Quality Technology* (Vol. 13, No. 2, 1981, pp. 111-114) describes and experiment that investigates the effects of four bleaching chemicals on pulp brightness. These four chemicals were selected at random from a large population of potential bleaching agents. The data are as follows:

Chemicals	Brightness				
1	77.199	74.466	92.746	76.208	82.876
2	80.522	79.306	81.914	80.346	73.385
3	79.417	78.017	91.596	80.802	80.626
4	78.001	78.358	77.544	77.364	77.386

(a) Is there a difference in the chemical types? Use Use α=0.05.

From the analysis below, there does not appear to be a difference in chemical types.

JMP Output

Summary of Fit

RSquare	0.123254
RSquare Adj	-0.04114
Root Mean Square Error	4.898921
Mean of Response	79.90395
Observations (or Sum Wgts)	20

Analysis of Variance

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Ratio**

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Ratio**
Model	3	53.98207	17.9940	0.7498
Error	16	383.99085	23.9994	**Prob > F**
C. Total	19	437.97292		0.5383

Effect Tests

Source	**Nparm**	**DF**	**Sum of Squares**	**F Ratio**	**Prob > F**
Chemical	3	3	53.982073	0.7498	0.5383

(b) Estimate the variability due to chemical types.

The JMP REML output below identifies the variance component for chemical types as -1.201081. This negative value is a concern. One solution would be to convert this to zero, but this has concerns as identified in Section 3.9.3 of the textbook. Another course of action is to re-estimate this variance component using a method that always provides a non-negative value. Another alternative is to assume that the underlying model is non-linear and re-examine the problem.

JMP Output

Parameter Estimates

Term	**Estimate**	**Std Error**	**DFDen**	**t Ratio**	**Prob>|t|**
Intercept	79.90395	0.948526	3	84.24	<.0001*

REML Variance Component Estimates

Random Effect	**Var Ratio**	**Var Component**	**Std Error**	**95% Lower**	**95% Upper**	**Pct of Total**
Chemical	-0.050046	-1.201081	3.3932473	-7.851723	5.4495617	-5.268
Residual		23.999428	8.4850792	13.312053	55.589101	105.268
Total		22.798347				100.000

Covariance Matrix of Variance Component Estimates

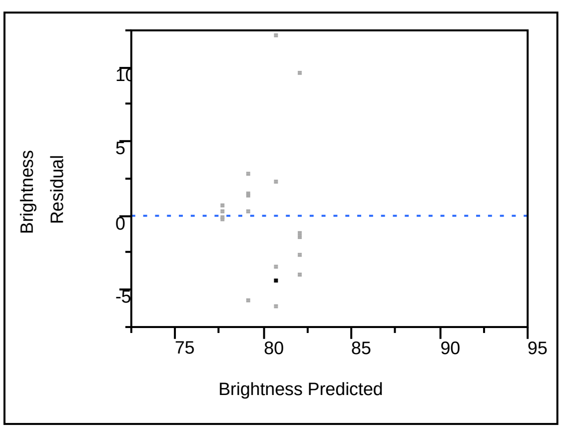
Random Effect	**Chemical**	**Residual**
Chemical	11.514127	-14.39931
Residual	-14.39931	71.996569

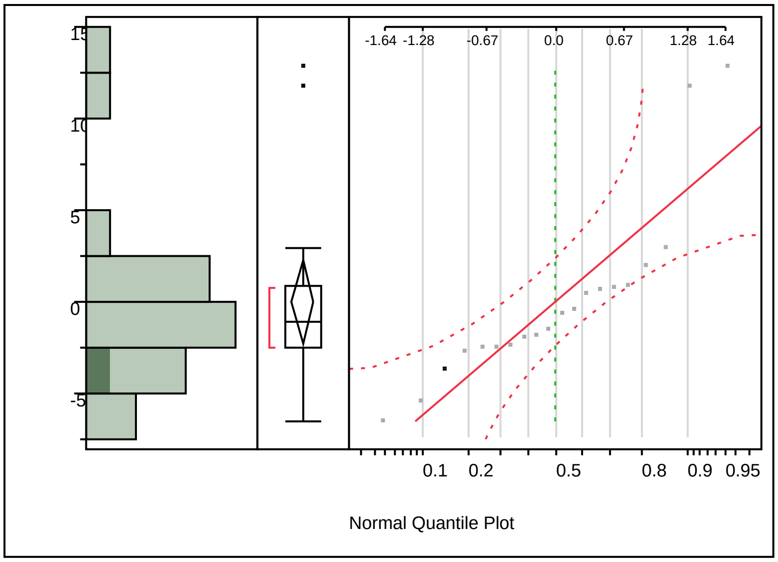
(c) Estimate the variability due to random error.

From the JMP REML output shown above, the variance component due to random error is 23.999428.

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

Examination of the residuals identified two outliers. These outliers correspond to the Brightness values of 92.746 and 91.596. The experimenter should resolve these outliers.





3.35. Consider the single-factor random effects model discussed in this chapter. Develop a procedure for finding a 100(1-α)% confidence interval on the ratio $\sigma_\tau^2 / \left(\sigma_\tau^2 + \sigma^2\right)$. Assume that the experiment is balanced.

The procedure shown below is based on the guidelines presented in Section 1.4 of the textbook. Rather than repeat the details of the seven steps, only additional information is provided below that is specific to the single-factor random effects and the confidence interval.

1. Recognition of and statement of the problem.
2. Selection of the response variable.
3. Choice of factors, levels, and range. For this case, one factor is chosen. However, the number of levels chosen and the number of replicates determines the degrees of freedom for the F value used in the confidence interval calculations. Because the levels are random, it is important to choose an adequate representation of this effect.
4. Choice of experimental design. For this case, the experimental design is a single factor experiment. As mentioned above, the number of replicates is important in the estimation of the confidence intervals. The value for α should also be identified as this could influence the number of replicates chosen.
5. Performing the experiment.
6. Statistical analysis of the data. Perform the analysis of variance in the same manner as a fixed effects case. Identify the MS_τ and MS_E from the ANOVA. Select the $F_{\alpha/2,a-1,N-a}$ and $F_{1-\alpha/2,a-1,N-a}$. Perform the calculations as identified in Equations 3.59a, 3,59b, and 3.60.
7. Conclusions and recommendations.

3.36. Consider testing the equality of the means of two normal populations, where the variances are unknown but are assumed to be equal. The appropriate test procedure is the pooled *t* test. Show that the pooled *t* test is equivalent to the single factor analysis of variance.

$$t_0 = \frac{\bar{y}_{1.} - \bar{y}_{2.}}{S_p\sqrt{\frac{2}{n}}} \sim t_{2n-2} \quad \text{assuming } n_1 = n_2 = n$$

$$S_p = \frac{\sum_{j=1}^{n}\left(y_{1j}-\bar{y}_{1.}\right)^2 + \sum_{j=1}^{n}\left(y_{2j}-\bar{y}_{2.}\right)^2}{2n-2} = \frac{\sum_{i=1}^{2}\sum_{j=1}^{n}\left(y_{ij}-\bar{y}_{1.}\right)^2}{2n-2} \equiv MS_E \quad \text{for a=2}$$

Furthermore, $\left(\bar{y}_{1.}-\bar{y}_{2.}\right)^2\left(\frac{n}{2}\right) = \sum_{i=1}^{2}\frac{y_{i.}^2}{n} - \frac{y_{..}^2}{2n}$, which is exactly the same as $SS_{\text{Treatments}}$ in a one-way classification with a=2. Thus we have shown that $t_0^2 = \frac{SS_{Treatments}}{MS_E}$. In general, we know that $t_u^2 = F_{1,u}$ so that $t_0^2 \sim F_{1,2n-2}$. Thus the square of the test statistic from the pooled *t*-test is the same test statistic that results from a single-factor analysis of variance with a=2.

3.37. Show that the variance of the linear combination $\sum_{i=1}^{a} c_i y_{i.}$ is $\sigma^2 \sum_{i=1}^{a} n_i c_i^2$.

$$V\left[\sum_{i=1}^{a} c_i y_{i.}\right] = \sum_{i=1}^{a} V\left(c_i y_{i.}\right) = \sum_{i=1}^{a} c_i^2 V\left[\sum_{j=1}^{n_i} y_{ij}\right] = \sum_{i=1}^{a} c_i^2 \sum_{j=1}^{n_i} V\left(y_{ij.}\right),\ V\left(y_{ij}\right) = \sigma^2$$

$$= \sum_{i=1}^{a} c_i^2 n_i \sigma^2$$

3.38. In a fixed effects experiment, suppose that there are *n* observations for each of four treatments. Let Q_1^2, Q_2^2, Q_3^2 be single-degree-of-freedom components for the orthogonal contrasts. Prove that $SS_{Treatments} = Q_1^2 + Q_2^2 + Q_3^2$.

$$C_1 = 3y_{1.} - y_{2.} - y_{3.} - y_{4.} \qquad SS_{C1} = Q_1^2$$
$$C_2 = 2y_{2.} - y_{3.} - y_{4.} \qquad SS_{C2} = Q_2^2$$
$$C_3 = y_{3.} - y_{4.} \qquad SS_{C3} = Q_3^2$$

$$Q_1^2 = \frac{(3y_{1.} - y_{2.} - y_{3.} - y_{4.})^2}{12n}$$

$$Q_2^2 = \frac{(2y_{2.} - y_{3.} - y_{4.})^2}{6n}$$

$$Q_3^2 = \frac{(y_{3.} - y_{4.})^2}{2n}$$

$$Q_1^2 + Q_2^2 + Q_3^2 = \frac{9\sum_{i=1}^{4} y_{i.}^2 - 6\left(\sum\sum_{i<j} y_{i.}y_{j.}\right)}{12n} \text{ and since}$$

$$\sum\sum_{i<j} y_{i.}y_{j.} = \frac{1}{2}\left(y_{..}^2 - \sum_{i=1}^{4} y_{i.}^2\right), \text{ we have } Q_1^2 + Q_2^2 + Q_3^2 = \frac{12\sum_{i=1}^{4} y_{i.}^2 - 3y_{..}^2}{12n} = \sum_{i=1}^{4}\frac{y_{i.}^2}{n} - \frac{y_{..}^2}{4n} = SS_{Treatments}$$

for a=4.

3.39. Use Bartlett's test to determine if the assumption of equal variances is satisfied in Problem 3.24. Use $\alpha = 0.05$. Did you reach the same conclusion regarding the equality of variance by examining the residual plots?

$$\chi_0^2 = 2.3026\frac{q}{c}\text{, where}$$

$$q = (N-a)log_{10}\,S_p^2 - \sum_{i=1}^{a}(n_i-1)log_{10}\,S_i^2$$

$$c = 1+\frac{1}{3(a-1)}\left(\sum_{i=1}^{a}(n_i-1)^{-1}-(N-a)^{-1}\right)$$

$$S_p^2 = \frac{\sum_{i=1}^{a}(n_i-1)S_i^2}{N-a}$$

$$S_1^2 = 11.2 \qquad S_2^2 = 14.8 \qquad S_3^2 = 20.8$$

$$S_p^2 = \frac{(5-1)11.2+(5-1)14.8+(5-1)20.8}{15-3} = 15.6$$

$$c = 1+\frac{1}{3(3-1)}\left(\sum_{i=1}^{a}(5-1)^{-1}-(15-3)^{-1}\right)$$

$$c = 1+\frac{1}{3(3-1)}\left(\frac{3}{4}+\frac{1}{12}\right) = 1.1389$$

$$q = (N-a)\log_{10}S_p^2 - \sum_{i=1}^{a}(n_i-1)\log_{10}S_i^2$$

$$q = (15-3)\log_{10}15.6 - 4\left(\log_{10}11.2+\log_{10}14.8+\log_{10}20.8\right)$$

$$q = 14.3175-14.150 = 0.1673$$

$$\chi_0^2 = 2.3026\frac{q}{c} = 2.3026\frac{0.1673}{1.1389} = 0.3382 \qquad \chi_{0.05,2}^2 = 5.99$$

Cannot reject null hypothesis; conclude that the variance are equal. This agrees with the residual plots in Problem 3.24.

3.40. Use the modified Levene test to determine if the assumption of equal variances is satisfied on Problem 3.24. Use $\alpha = 0.05$. Did you reach the same conclusion regarding the equality of variances by examining the residual plots?

The absolute value of Battery Life – brand median is:

	$\lvert y_{ij} - \boxed{y_i}\rvert$	
Brand 1	Brand 2	Brand 3
4	4	8
0	0	0
4	5	4
0	4	2
4	2	0

The analysis of variance indicates that there is not a difference between the different brands and therefore the assumption of equal variances is satisfied.

Design Expert Output

Response: Mod Levine
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	0.93	2	0.47	0.070	0.9328
A	*0.93*	*2*	*0.47*	*0.070*	*0.9328*
Pure Error	80.00	12	6.67		
Cor Total	80.93	14			

3.41. Refer to Problem 3.22. If we wish to detect a maximum difference in mean response times of 10 milliseconds with a probability of at least 0.90, what sample size should be used? How would you obtain a preliminary estimate of σ^2?

$$\Phi^2 = \frac{nD^2}{2a\sigma^2}\text{, use } MS_E \text{ from Problem 3.20 to estimate } \sigma^2.$$

$$\Phi^2 = \frac{n(10)^2}{2(3)(16.9)} = 0.986n$$

Letting $\alpha = 0.05$, P(accept) = 0.1 , $\upsilon_1 = a - 1 = 2$

Trial and Error yields:

n	υ_2	Φ	P(accept)
5	12	2.22	0.17
6	15	2.43	0.09
7	18	2.62	0.04

Choose n ≥ 6, therefore N ≥ 18

Notice that we have used an estimate of the variance obtained from the present experiment. This indicates that we probably didn't use a large enough sample (*n* was 5 in problem 3.20) to satisfy the criteria specified in this problem. However, the sample size *was* adequate to detect differences in one of the circuit types.

When we have no prior estimate of variability, sometimes we will generate sample sizes for a range of possible variances to see what effect this has on the size of the experiment. Often a knowledgeable expert will be able to bound the variability in the response, by statements such as "the standard deviation is going to be *at least*…" or "the standard deviation shouldn't be larger than…".

3.42. Refer to Problem 3.26.

(a) If we wish to detect a maximum difference in mean battery life of 10 hours with a probability of at least 0.90, what sample size should be used? Discuss how you would obtain a preliminary estimate of σ^2 for answering this question.

Use the MS_E from Problem 3.26.

$$\Phi^2 = \frac{nD^2}{2a\sigma^2} \qquad \Phi^2 = \frac{n(10)^2}{2(3)(15.60)} = 1.0684n$$

Letting $\alpha = 0.05$, P(accept) = 0.1 , $\upsilon_1 = a - 1 = 2$

Trial and Error yields:

n	υ_2	Φ	*P(accept)*
4	9	2.067	0.25
5	12	2.311	0.12
6	15	2.532	0.05

Choose n ≥ 6, therefore N ≥ 18

See the discussion from the previous problem about the estimate of variance.

(b) If the difference between brands is great enough so that the standard deviation of an observation is increased by 25 percent, what sample size should be used if we wish to detect this with a probability of at least 0.90?

$$\upsilon_1 = a - 1 = 2 \quad \upsilon_2 = N - a = 15 - 3 = 12 \quad \alpha = 0.05 \quad P(\,accept\,) \le 0.1$$
$$\lambda = \sqrt{1 + n\left[(1 + 0.01P)^2 - 1\right]} = \sqrt{1 + n\left[(1 + 0.01(25))^2 - 1\right]} = \sqrt{1 + 0.5625n}$$

Trial and Error yields:

n	υ_2	Φ	*P(accept)*
8	21	2.12	0.16
9	24	2.25	0.13
10	27	2.37	0.09

Choose n ≥ 10, therefore N ≥ 30

3.43. Consider the experiment in Problem 3.26. If we wish to construct a 95 percent confidence interval on the difference in two mean battery lives that has an accuracy of ±2 weeks, how many batteries of each brand must be tested?

$$\alpha = 0.05 \qquad MS_E = 15.6$$
$$width = t_{0.025,N-a}\sqrt{\frac{2MS_E}{n}}$$

Trial and Error yields:

n	υ_2	*t*	width
5	12	2.179	5.44
10	27	2.05	3.62
31	90	1.99	1.996
32	93	1.99	1.96

Choose n ≥ 31, therefore N ≥ 93

3.44. Suppose that four normal populations have means of μ_1=50, μ_2=60, μ_3=50, and μ_4=60. How many observations should be taken from each population so that the probability or rejecting the null hypothesis of equal population means is at least 0.90? Assume that α=0.05 and that a reasonable estimate of the error variance is σ^2=25.

$$\mu_i = \mu + \tau_i , i = 1,2,3,4$$
$$\mu = \frac{\sum_{i=1}^{4} \mu_i}{4} = \frac{220}{4} = 55$$
$$\tau_1 = -5, \tau_2 = 5, \tau_3 = -5, \tau_4 = 5$$
$$\sum_{i=1}^{4} \tau_i^2 = 100$$

$$\Phi^2 = \frac{n\sum \tau_i^2}{a\sigma^2} = \frac{100n}{4(25)} = n$$
$$\Phi = \sqrt{n}$$

$\upsilon_1 = 3, \upsilon_2 = 4(n-1), \alpha = 0.05$, From the O.C. curves we can construct the following:

n	Φ	υ_2	β	1-β
4	2.00	12	0.18	0.82
5	2.24	16	0.08	0.92

Therefore, select n=5

3.45. Refer to Problem 3.44.

(a) How would your answer change if a reasonable estimate of the experimental error variance were σ^2 = 36?

$$\Phi^2 = \frac{n\sum \tau_i^2}{a\sigma^2} = \frac{100n}{4(36)} = 0.6944n$$
$$\Phi = \sqrt{0.6944n}$$

$\upsilon_1 = 3, \upsilon_2 = 4(n-1), \alpha = 0.05$, From the O.C. curves we can construct the following:

n	Φ	υ_2	β	1-β
5	1.863	16	0.24	0.76
6	2.041	20	0.15	0.85
7	2.205	24	0.09	0.91

Therefore, select n=7

(b) How would your answer change if a reasonable estimate of the experimental error variance were σ^2 = 49?

$$\Phi^2 = \frac{n\sum \tau_i^2}{a\sigma^2} = \frac{100n}{4(49)} = 0.5102n$$
$$\Phi = \sqrt{0.5102n}$$

$\upsilon_1 = 3, \upsilon_2 = 4(n-1), \alpha = 0.05$, From the O.C. curves we can construct the following:

n	Φ	υ_2	β	1-β
7	1.890	24	0.16	0.84
8	2.020	28	0.11	0.89
9	2.142	32	0.09	0.91

Therefore, select n=9

(c) Can you draw any conclusions about the sensitivity of your answer in the particular situation about how your estimate of σ affects the decision about sample size?

As our estimate of variability increases the sample size must increase to ensure the same power of the test.

(d) Can you make any recommendations about how we should use this general approach to choosing *n* in practice?

When we have no prior estimate of variability, sometimes we will generate sample sizes for a range of possible variances to see what effect this has on the size of the experiment. Often a knowledgeable expert will be able to bound the variability in the response, by statements such as "the standard deviation is going to be *at least*…" or "the standard deviation shouldn't be larger than…".

3.46. Refer to the aluminum smelting experiment described in Section 3.8.3. Verify that ratio control methods do not affect average cell voltage. Construct a normal probability plot of residuals. Plot the residuals versus the predicted values. Is there an indication that any underlying assumptions are violated?

Design Expert Output

Response: Cell 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	2.746E-003	3	9.153E-004	0.20	0.8922	not significant
A	*2.746E-003*	*3*	*9.153E-004*	*0.20*	*0.8922*	
Residual	0.090	20	4.481E-003			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.090*	*20*	*4.481E-003*			
Cor Total	0.092	23				

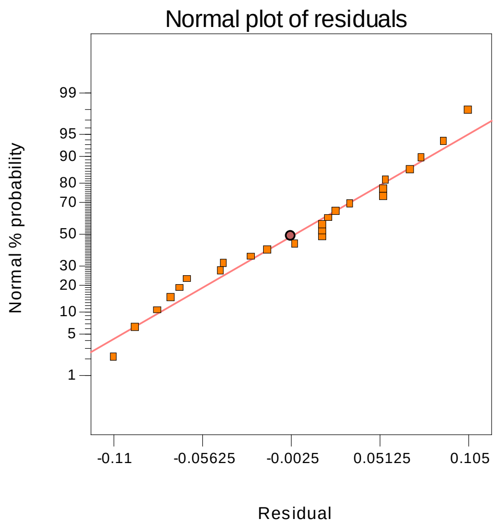
The "Model F-value" of 0.20 implies the model is not significant relative to the noise. There is a 89.22 % chance that a "Model F-value" this large could occur due to noise.

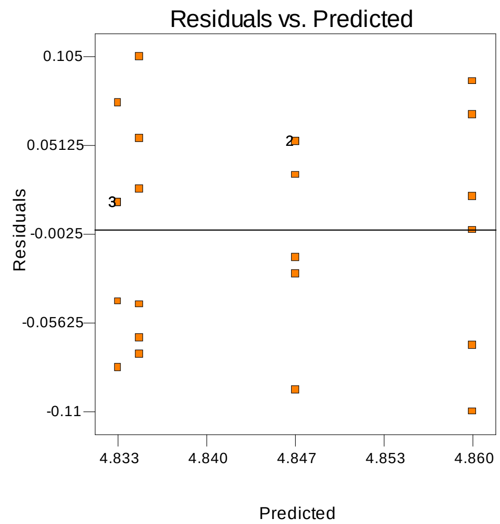
Treatment Means (Adjusted, If Necessary)

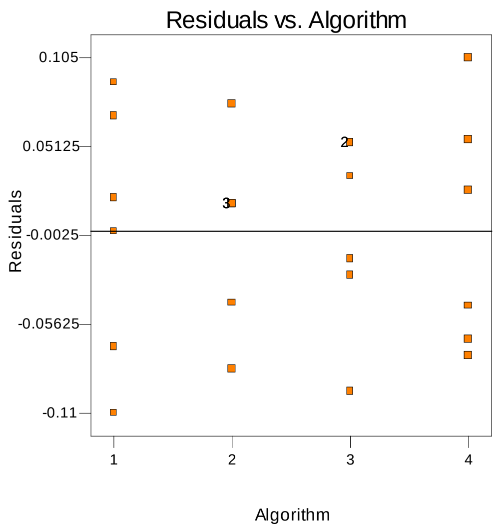
	Estimated Mean	Standard Error
1-1	4.86	0.027
2-2	4.83	0.027
3-3	4.85	0.027
4-4	4.84	0.027

Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	0.027	1	0.039	0.69	0.4981
1 vs 3	0.013	1	0.039	0.35	0.7337
1 vs 4	0.025	1	0.039	0.65	0.5251
2 vs 3	-0.013	1	0.039	-0.35	0.7337
2 vs 4	-1.667E-003	1	0.039	-0.043	0.9660
3 vs 4	0.012	1	0.039	0.30	0.7659

The following residual plots are satisfactory.







3.47. Refer to the aluminum smelting experiment in Section 3.8.3. Verify the ANOVA for pot noise summarized in Table 3.16. Examine the usual residual plots and comment on the experimental validity.

Design Expert Output

Response: Cell StDev 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	6.17	3	2.06	21.96	< 0.0001	significant
A	*6.17*	*3*	*2.06*	*21.96*	*< 0.0001*	
Residual	1.87	20	0.094			
Lack of Fit	*0.000*	*0*				
Pure Error	*1.87*	*20*	*0.094*			
Cor Total	8.04	23				

The Model F-value of 21.96 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

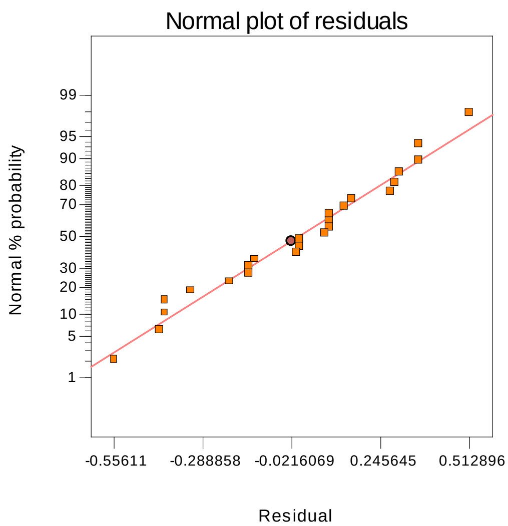
Treatment Means (Adjusted, If Necessary)

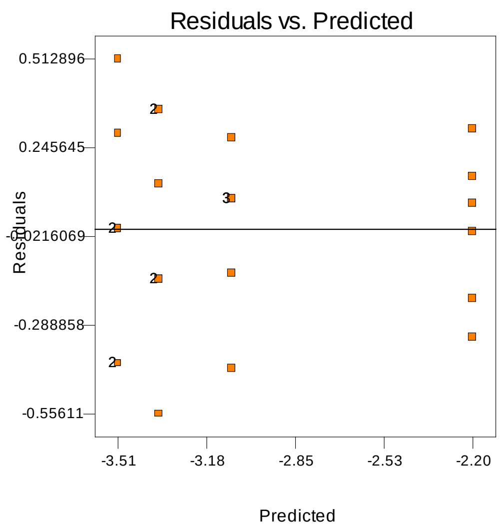
Estimated Standard

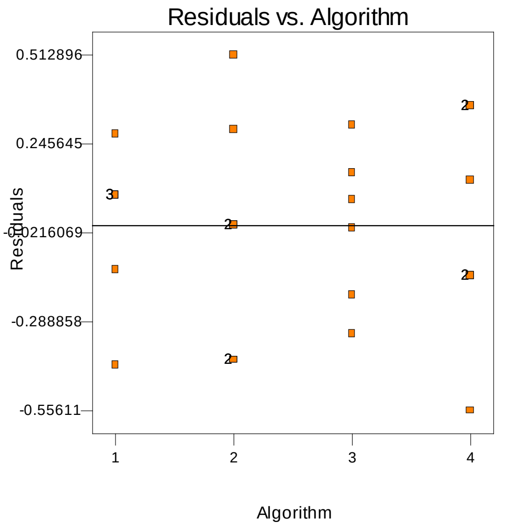
	Mean	Error
1-1	-3.09	0.12
2-2	-3.51	0.12
3-3	-2.20	0.12
4-4	-3.36	0.12

Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	0.42	1	0.18	2.38	0.0272
1 vs 3	-0.89	1	0.18	-5.03	< 0.0001
1 vs 4	0.27	1	0.18	1.52	0.1445
2 vs 3	-1.31	1	0.18	-7.41	< 0.0001
2 vs 4	-0.15	1	0.18	-0.86	0.3975
3 vs 4	1.16	1	0.18	6.55	< 0.0001

The following residual plots identify the residuals to be normally distributed, randomly distributed through the range of prediction, and uniformly distributed across the different algorithms. This validates the assumptions for the experiment.







3.48. Four different feed rates were investigated in an experiment on a CNC machine producing a component part used in an aircraft auxiliary power unit. The manufacturing engineer in charge of the experiment knows that a critical part dimension of interest may be affected by the feed rate. However, prior experience has indicated that only dispersion effects are likely to be present. That is, changing the feed rate does not affect the average dimension, but it could affect dimensional variability. The engineer makes five production runs at each feed rate and obtains the standard deviation of the critical dimension (in 10^{-3} mm). The data are shown below. Assume that all runs were made in random order.

Feed Rate		Production	Run		
(in/min)	1	2	3	4	5
10	0.09	0.10	0.13	0.08	0.07
12	0.06	0.09	0.12	0.07	0.12
14	0.11	0.08	0.08	0.05	0.06
16	0.19	0.13	0.15	0.20	0.11

(a) Does feed rate have any effect on the standard deviation of this critical dimension?

Because the residual plots were not acceptable for the non-transformed data, a square root transformation was applied to the standard deviations of the critical dimension. Based on the computer output below, the feed rate has an effect on the standard deviation of the critical dimension.

Design Expert Output

Response: Run StDev Transform: Square root 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	0.040	3	0.013	7.05	0.0031	significant
A	*0.040*	*3*	*0.013*	*7.05*	*0.0031*	
Residual	0.030	16	1.903E-003			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.030*	*16*	*1.903E-003*			
Cor Total	0.071	19				

The Model F-value of 7.05 implies the model is significant. There is only a 0.31% chance that a "Model F-Value" this large could occur due to noise.

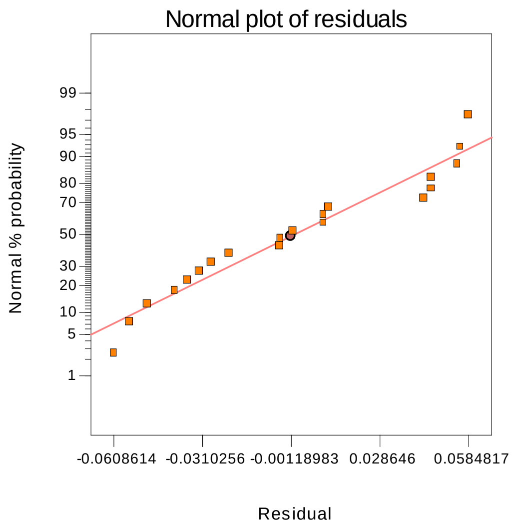
Treatment Means (Adjusted, If Necessary)

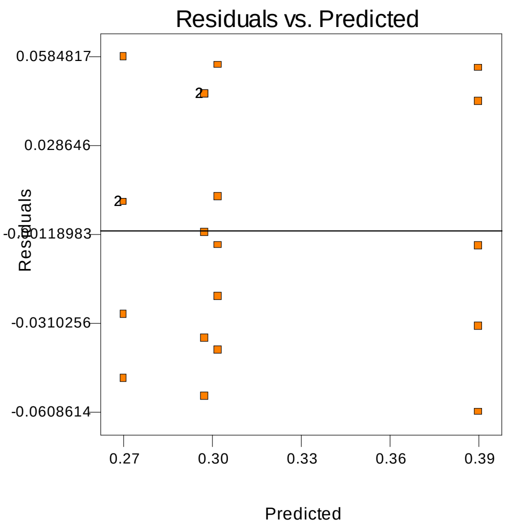
	Estimated Mean	Standard Error
1-10	0.30	0.020
2-12	0.30	0.020
3-14	0.27	0.020
4-16	0.39	0.020

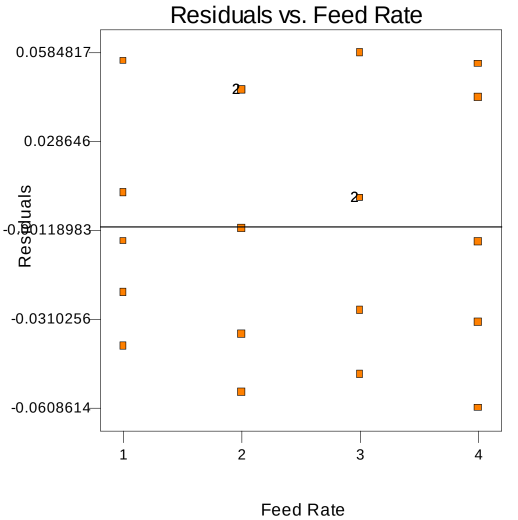
Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	4.371E-003	1	0.028	0.16	0.8761
1 vs 3	0.032	1	0.028	1.15	0.2680
1 vs 4	-0.088	1	0.028	-3.18	0.0058
2 vs 3	0.027	1	0.028	0.99	0.3373
2 vs 4	-0.092	1	0.028	-3.34	0.0042
3 vs 4	-0.12	1	0.028	-4.33	0.0005

(b) Use the residuals from this experiment of investigate model adequacy. Are there any problems with experimental validity?

The residual plots are satisfactory.







3.49. Consider the data shown in Problem 3.22.

(a) Write out the least squares normal equations for this problem, and solve them for $\hat{\mu}$ and $\hat{\tau}_i$, using the usual constraint $\left(\sum_{i=1}^{3}\hat{\tau}_i = 0\right)$. Estimate $\tau_1 - \tau_2$.

$$
\begin{array}{lllll}
15\hat{\mu} & +5\hat{\tau}_1 & +5\hat{\tau}_2 & +5\hat{\tau}_3 & =207 \\
5\hat{\mu} & +5\hat{\tau}_1 & & & =54 \\
5\hat{\mu} & & +5\hat{\tau}_2 & & =111 \\
5\hat{\mu} & & & +5\hat{\tau}_3 & =42
\end{array}
$$

Imposing $\sum_{i=1}^{3}\hat{\tau}_i = 0$, therefore $\hat{\mu} = 13.80$, $\hat{\tau}_1 = -3.00$, $\hat{\tau}_2 = 8.40$, $\hat{\tau}_3 = -5.40$

$$\hat{\tau}_1 - \hat{\tau}_2 = -3.00 - 8.40 = -11.40$$

(b) Solve the equations in (a) using the constraint $\hat{\tau}_3 = 0$. Are the estimators $\hat{\tau}_i$ and $\hat{\mu}$ the same as you found in (a)? Why? Now estimate $\tau_1 - \tau_2$ and compare your answer with that for (a). What statement can you make about estimating contrasts in the τ_i?

Imposing the constraint, $\hat{\tau}_3 = 0$ we get the following solution to the normal equations: $\hat{\mu} = 8.40$, $\hat{\tau}_1 = 2.40$, $\hat{\tau}_2 = 13.8$, and $\hat{\tau}_3 = 0$. These estimators are not the same as in part (a). However, $\hat{\tau}_1 - \hat{\tau}_2 = 2.40 - 13.80 = -11.40$, is the same as in part (a). The contrasts are estimable.

(c) Estimate $\mu + \tau_1$, $2\tau_1 - \tau_2 - \tau_3$ and $\mu + \tau_1 + \tau_2$ using the two solutions to the normal equations. Compare the results obtained in each case.

	Contrast	Estimated from Part (a)	Estimated from Part (b)
1	$\mu + \tau_1$	10.80	10.80
2	$2\tau_1 - \tau_2 - \tau_3$	-9.00	-9.00
3	$\mu + \tau_1 + \tau_2$	19.20	24.60

Contrasts 1 and 2 are estimable, 3 is not estimable.

3.50. Apply the general regression significance test to the experiment in Example 3.5. Show that the procedure yields the same results as the usual analysis of variance.

From the etch rate table:

$$y_{..} = 12355$$

from Example 3.5, we have:

$$\hat{\mu} = 617.75 \quad \hat{\tau}_1 = -66.55 \quad \hat{\tau}_2 = -30.35$$
$$\hat{\tau}_3 = 7.65 \quad \hat{\tau}_4 = 89.25$$

$\sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 = 7,704,511$, with 20 degrees of freedom.

$$\begin{aligned} R(\mu,\tau) &= \hat{\mu}y_{..} + \sum_{i=1}^{5}\hat{\tau}y_{i.} \\ &= (617.75)(12355) + (-66.55)(2756) + (-30.35)(2937) + (7.65)(3127) + (89.25)(3535) \\ &= 7,632,301.25 + 66,870.55 = 7,699,171.80 \end{aligned}$$

with 4 degrees of freedom.

$$SS_E = \sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 - R(\mu,\tau) = 7,704,511 - 7,699,171.80 = 5339.2$$

with 20-4 degrees of freedom.

This is identical to the SS_E found in Example 3.5.

The reduced model:

$R(\mu) = \hat{\mu} y_{..} = (617.75)(12355) = 7,632,301.25$, with 1 degree of freedom.

$R(\tau|\mu) = R(\mu,\tau) - R(\mu) = 7,699,171.80 - 7,632,301.25 = 66,870.55$, with 4-1=3 degrees of freedom.

Note: $R(\tau|\mu) = SS_{Treatment}$ from Example 3.1.

Finally,

$$F_0 = \frac{\frac{R(\tau|\mu)}{3}}{\frac{SS_E}{16}} = \frac{\frac{66,870.55}{3}}{\frac{5339.2}{16}} = \frac{22290.8}{333.7} = 66.8$$

which is the same as computed in Example 3.5.

3.51. Use the Kruskal-Wallis test for the experiment in Problem 3.23. Are the results comparable to those found by the usual analysis of variance?

From Design Expert Output of Problem 3.21

Response: Life in in h
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	30.17	3	10.06	3.05	0.0525	not significant
A	*30.16*	*3*	*10.05*	*3.05*	*0.0525*	
Residual	65.99	20	3.30			
Lack of Fit	*0.000*	*0*				
Pure Error	*65.99*	*20*	*3.30*			
Cor Total	96.16	23				

$$H = \frac{12}{N(N+1)}\left[\sum_{i=1}^{a}\frac{R_{i.}^2}{n_i}\right] - 3(N+1) = \frac{12}{24(24+1)}[4060.75] - 3(24+1) = 6.22$$

$$\chi^2_{0.05,3} = 7.81$$

Accept the null hypothesis; the treatments are not different. This agrees with the analysis of variance.

3.52. Use the Kruskal-Wallis test for the experiment in Problem 3.24. Compare conclusions obtained with those from the usual analysis of variance?

From Design Expert Output of Problem 3.22

Response: Noise
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	12042.00	3	4014.00	21.78	< 0.0001	significant
A	*12042.00*	*3*	*4014.00*	*21.78*	*< 0.0001*	
Residual	2948.80	16	184.30			
Lack of Fit	*0.000*	*0*				
Pure Error	*2948.80*	*16*	*184.30*			
Cor Total	14990.80	19				

$$H = \frac{12}{N(N+1)}\left[\sum_{i=1}^{a}\frac{R_{i.}^2}{n_i}\right] - 3(N+1) = \frac{12}{20(20+1)}[2726.8] - 3(20+1) = 14.91$$

$$\chi^2_{0.05,3} = 7.81$$

Reject the null hypothesis because the treatments are different. This agrees with the analysis of variance.

3.53. Consider the experiment in Example 3.5. Suppose that the largest observation on etch rate is incorrectly recorded as 250A/min. What effect does this have on the usual analysis of variance? What effect does it have on the Kruskal-Wallis test?

The incorrect observation reduces the analysis of variance F_0 from 66.8 to 0.50. It does change the value of the Kruskal-Wallis test statistic but not the result.

Minitab Output

One-way ANOVA: Etch Rate 2 versus Power

```
Analysis of Variance for Etch Rat
Source     DF        SS        MS        F        P
Power       3     15927      5309     0.50    0.685
Error      16    168739     10546
Total      19    184666
```

3.54 A textile mill has a large number of looms. Each loom is supposed to provide the same output of cloth per minute. To investigate this assumption, five looms are chosen at random, and their output is noted at different times. The following data are obtained:

Loom	Output				
1	14.0	14.1	14.2	14.0	14.1
2	13.9	13.8	13.9	14.0	14.0
3	14.1	14.2	14.1	14.0	13.9
4	13.6	13.8	14.0	13.9	13.7
5	13.8	13.6	13.9	13.8	14.0

(a) Explain why this is a random effects experiment. Are the looms equal in output? Use α=0.05.

(b) Estimate the variability between looms.

(c) Estimate the experimental error variance.

(d) Find a 95 percent confidence interval for sdfdsfseererrw(need equation typed here)

(e) Analyze the residuals from this experiment. Do you think that the analysis of variance assumptions are satisfied?

(f) Use the REML method to analyze this data. Compare the 95 percent confidence interval on the error variance from REML with the exact chi-square confidence interval.

3.55 A manufacturer suspects that the batches of raw material furnished by his supplier differ significantly in calcium content. There are a large number of batches currently in the warehouse. Five of these are randomly selected for study. A chemist makes five determinations on each batch as obtains the following data:

Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
23.46	23.59	23.51	23.28	23.29
23.48	23.46	23.64	23.40	23.46
23.56	23.42	23.46	23.37	23.37
23.39	23.49	23.52	23.46	23.32
23.40	23.50	23.49	23.39	23.38

This is the same as question 3.30 except for (e).

(a) Is there significant variation in the calcium content from batch to batch? Use α=0.05.

(b) Estimate the components of variance.

(c) Find a 95 percent confidence interval for sdfdsfseererrw(need equation typed here)

(d) Analyze the residuals from this experiment. Are the analysis of variance assumptions are satisfied?

(e) Use the REML method to analyze this data. Compare the 95 percent confidence interval on the error variance from REML with the exact chi-square confidence interval.

Chapter 4
Randomized Blocks, Latin Squares, and Related Designs
Solutions

4.1. The ANOVA from a randomized complete block experiment output is shown below.

Source	*DF*	*SS*	*MS*	*F*	*P*
Treatment	4	1010.56	?	29.84	?
Block	?	?	64.765	?	?
Error	20	169.33	?		
Total	29	1503.71			

(a) Fill in the blanks. You may give bounds on the *P*-value.

Completed table is:

Source	*DF*	*SS*	*MS*	*F*	*P*
Treatment	4	1010.56	252.640	29.84	< 0.00001
Block	5	323.82	64.765		
Error	20	169.33	8.467		
Total	29	1503.71			

(b) How many blocks were used in this experiment?

Six blocks were used.

(c) What conclusions can you draw?

The treatment effect is significant; the means of the five treatments are not all equal.

4.2. Consider the single-factor completely randomized experiment shown in Problem 3.4. Suppose that this experiment had been conducted in a randomized complete block design, and that the sum of squares for blocks was 80.00. Modify the ANOVA for this experiment to show the correct analysis for the randomized complete block experiment.

The modified ANOVA is shown below:

Source	DF	SS	MS	F	P
Treatment	4	987.71	246.93	46.3583	< 0.00001
Block	5	80.00	16.00		
Error	20	106.53	5.33		
Total	29	1174.24			

4.3. A chemist wishes to test the effect of four chemical agents on the strength of a particular type of cloth. Because there might be variability from one bolt to another, the chemist decides to use a randomized block design, with the bolts of cloth considered as blocks. She selects five bolts and applies all four chemicals in random order to each bolt. The resulting tensile strengths follow. Analyze the data from this experiment (use $\alpha = 0.05$) and draw appropriate conclusions.

	Bolt				
Chemical	1	2	3	4	5
1	73	68	74	71	67
2	73	67	75	72	70
3	75	68	78	73	68
4	73	71	75	75	69

Design Expert Output

Response: Strength

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	157.00	4	39.25			
Model	12.95	3	4.32	2.38	0.1211	not significant
A	*12.95*	*3*	*4.32*	*2.38*	*0.1211*	
Residual	21.80	12	1.82			
Cor Total	191.75	19				

The "Model F-value" of 2.38 implies the model is not significant relative to the noise. There is a 12.11 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	1.35	R-Squared	0.3727
Mean	71.75	Adj R-Squared	0.2158
C.V.	1.88	Pred R-Squared	-0.7426
PRESS	60.56	Adeq Precision	10.558

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	70.60	0.60
2-2	71.40	0.60
3-3	72.40	0.60
4-4	72.60	0.60

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-0.80	1	0.85	-0.94	0.3665
1 vs 3	-1.80	1	0.85	-2.11	0.0564
1 vs 4	-2.00	1	0.85	-2.35	0.0370
2 vs 3	-1.00	1	0.85	-1.17	0.2635
2 vs 4	-1.20	1	0.85	-1.41	0.1846
3 vs 4	-0.20	1	0.85	-0.23	0.8185

There is no difference among the chemical types at $\alpha = 0.05$ level.

4.4. Three different washing solutions are being compared to study their effectiveness in retarding bacteria growth in five-gallon milk containers. The analysis is done in a laboratory, and only three trials can be run on any day. Because days could represent a potential source of variability, the experimenter decides to use a randomized block design. Observations are taken for four days, and the data are shown here. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

	Days			
Solution	1	2	3	4
1	13	22	18	39
2	16	24	17	44
3	5	4	1	22

Design Expert Output

Response: 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	
Block	1106.92	3	368.97			
Model	703.50	2	351.75	40.72	0.0003	significant
A	*703.50*	*2*	*351.75*	*40.72*	*0.0003*	
Residual	51.83	6	8.64			
Cor Total	1862.25	11				

The Model F-value of 40.72 implies the model is significant. There is only a 0.03% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.94	R-Squared	0.9314
Mean	18.75	Adj R-Squared	0.9085
C.V.	15.68	Pred R-Squared	0.7255
PRESS	207.33	Adeq Precision	19.687

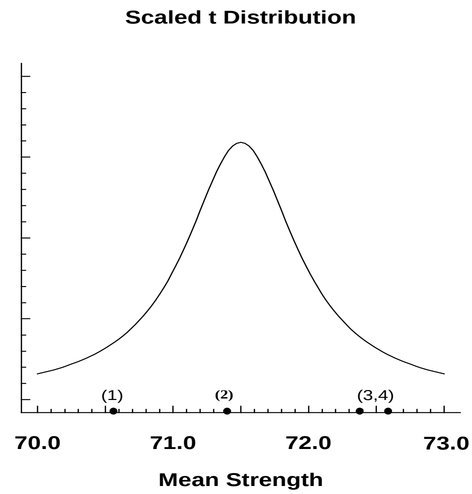
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	23.00	1.47
2-2	25.25	1.47
3-3	8.00	1.47

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-2.25	1	2.08	-1.08	0.3206
1 vs 3	15.00	1	2.08	7.22	0.0004
2 vs 3	17.25	1	2.08	8.30	0.0002

There is a difference between the means of the three solutions. The Fisher LSD procedure indicates that solution 3 is significantly different than the other two.

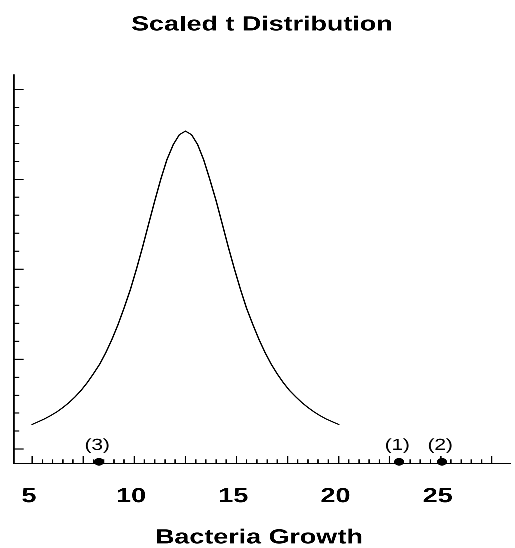
4.5. Plot the mean tensile strengths observed for each chemical type in Problem 4.3 and compare them to a scaled *t* distribution. What conclusions would you draw from the display?



$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{b}} = \sqrt{\frac{1.82}{5}} = 0.603$$

There is no obvious difference between the means. This is the same conclusion given by the analysis of variance.

4.6. Plot the average bacteria counts for each solution in Problem 4.4 and compare them to an appropriately scaled *t* distribution. What conclusions can you draw?



$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{b}} = \sqrt{\frac{8.64}{4}} = 1.47$$

There is no difference in mean bacteria growth between solutions 1 and 2. However, solution 3 produces significantly lower mean bacteria growth. This is the same conclusion reached from the Fisher LSD procedure in Problem 4.4.

4.7. Consider the hardness testing experiment described in Section 4.1. Suppose that the experiment was conducted as described and the following Rockwell C-scale data (coded by subtracting 40 units) obtained:

	Coupon			
Tip	1	2	3	4
1	9.3	9.4	9.6	10.0
2	9.4	9.3	9.8	9.9
3	9.2	9.4	9.5	9.7
4	9.7	9.6	10.0	10.2

(a) Analyize the data from this experiment.

There is a difference between the means of the four tips.

Design Expert Output

Response: Hardness

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**	
Bock	0.82	3	0.27			
Model	0.38	3	0.13	14.44	0.0009	significant
A	0.38	3	0.13	14.44	0.0009	
Residual	0.080	9	8.889E-003			
Cor Total	1.29	15				

The Model F-value of 14.44 implies the model is significant. There is only a 0.09% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.094	R-Squared	0.8280
Mean	9.63	Adj R-Squared	0.7706
C.V.	0.98	Pred R-Squared	0.4563
PRESS	0.25	Adeq Precision	15.635

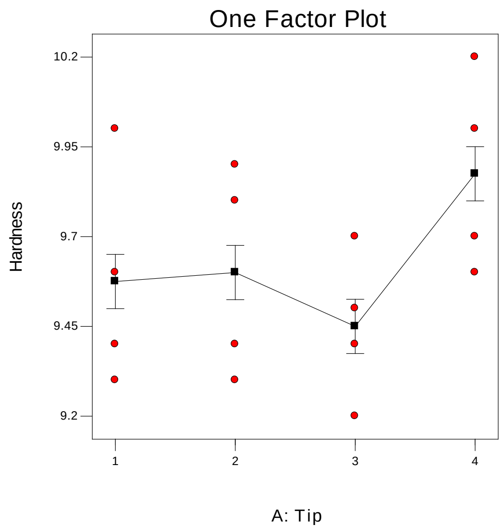
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	9.57	0.047
2-2	9.60	0.047
3-3	9.45	0.047
4-4	9.88	0.047

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-0.025	1	0.067	-0.38	0.7163
1 vs 3	0.13	1	0.067	1.87	0.0935
1 vs 4	-0.30	1	0.067	-4.50	0.0015
2 vs 3	0.15	1	0.067	2.25	0.0510
2 vs 4	-0.27	1	0.067	-4.12	0.0026
3 vs 4	-0.43	1	0.067	-6.37	0.0001

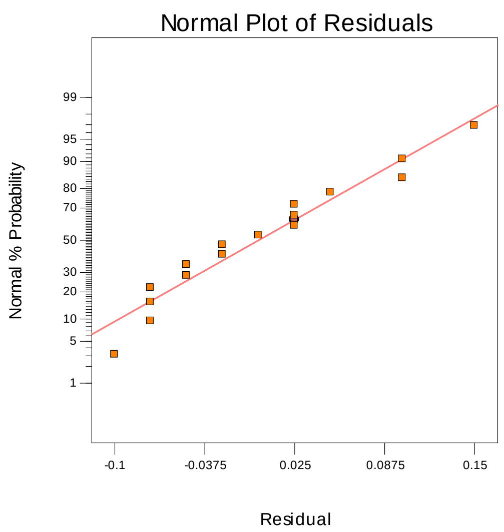
(b) Use the Fisher LSD method to make comparisons among the four tips to determine specifically which tips differ in mean hardness readings.

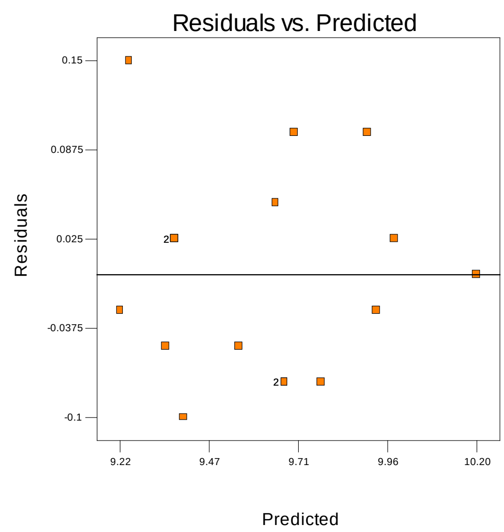
Based on the LSD bars in the Design Expert plot below, the mean of tip 4 differs from the means of tips 1, 2, and 3. The LSD method identifies a marginal difference between the means of tips 2 and 3.

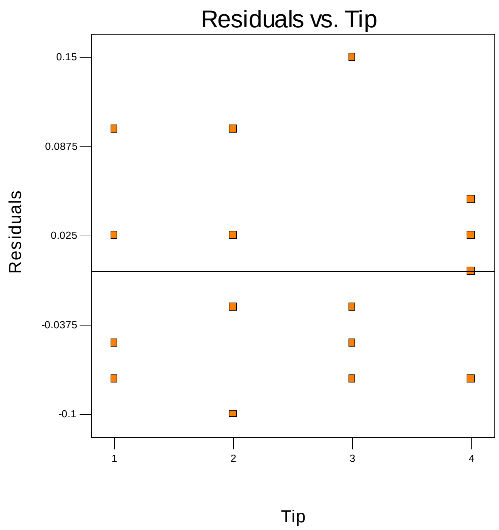


(c) Analyze the residuals from this experiment.

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







4.8. A consumer products company relies on direct mail marketing pieces as a major component of its advertising campaigns. The company has three different designs for a new brochure and want to evaluate their effectiveness, as there are substantial differences in costs between the three designs. The company decides to test the three designs by mailing 5,000 samples of each to potential customers in four different regions of the country. Since there are known regional differences in the customer base, regions are considered as blocks. The number of responses to each mailing is shown below.

	Region			
Design	NE	NW	SE	SW
1	250	350	219	375
2	400	525	390	580
3	275	340	200	310

(a) Analyze the data from this experiment.

The residuals of the analsysis below identify concerns with the normality and equality of variance assumptions. As a result, a square root transformation was applied as shown in the second ANOVA table. The residuals of both analysis are presented for comparison in part (c) of this problem. The analysis concludes that there is a difference between the mean number of responses for the three designs.

Design Expert Output

Response: Number of responses
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**	
Block	49035.67	3	16345.22			
Model	90755.17	2	45377.58	50.15	0.0002	significant
A	*90755.17*	*2*	*45377.58*	*50.15*	*0.0002*	
Residual	5428.83	6	904.81			
Cor Total	1.452E+005	11				

The Model F-value of 50.15 implies the model is significant. There is only a 0.02% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	30.08	R-Squared	0.9436
Mean	351.17	Adj R-Squared	0.9247
C.V.	8.57	Pred R-Squared	0.7742
PRESS	21715.33	Adeq Precision	16.197

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	298.50	15.04
2-2	473.75	15.04
3-3	281.25	15.04

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H_0 Coeff=0**	**Prob > |t|**
1 vs 2	-175.25	1	21.27	-8.24	0.0002
1 vs 3	17.25	1	21.27	0.81	0.4483
2 vs 3	192.50	1	21.27	9.05	0.0001

Design Expert Output for Transformed Data

Response: Number of responses 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**	
Block	35.89	3	11.96			
Model	60.73	2	30.37	60.47	0.0001	significant
A	*60.73*	*2*	*30.37*	*60.47*	*0.0001*	
Residual	3.01	6	0.50			
Cor Total	99.64	11				

The Model F-value of 60.47 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.71	R-Squared	0.9527
Mean	18.52	Adj R-Squared	0.9370
C.V.	3.83	Pred R-Squared	0.8109
PRESS	12.05	Adeq Precision	18.191

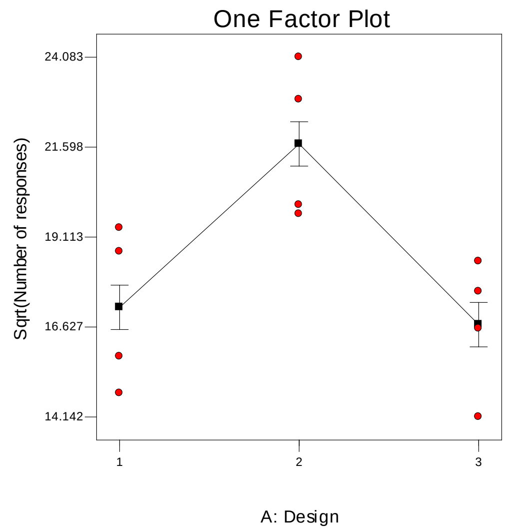
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-1	17.17	0.35
2-2	21.69	0.35
3-3	16.69	0.35

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H_0 Coeff=0**	**Prob > |t|**
1 vs 2	-4.52	1	0.50	-9.01	0.0001
1 vs 3	0.48	1	0.50	0.95	0.3769
2 vs 3	4.99	1	0.50	9.96	< 0.0001

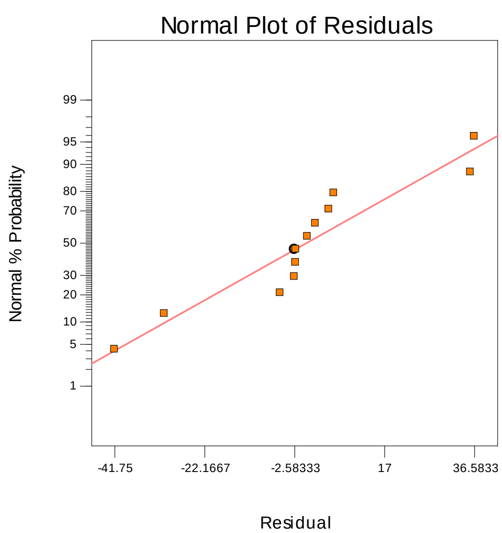
(b) Use the Fisher LSD method to make comparisons among the three designs to determine specifically which designs differ in mean response rate.

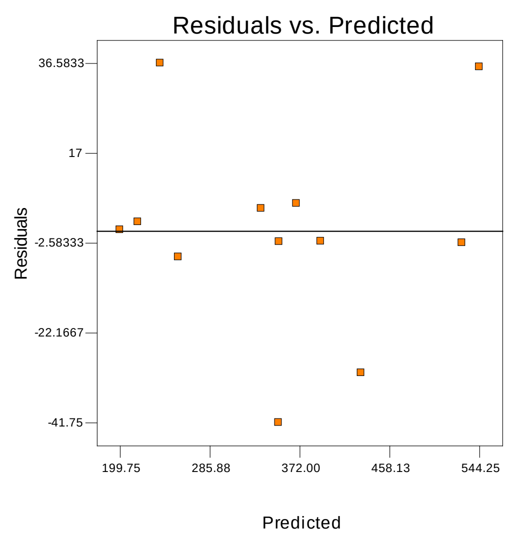
Based on the LSD bars in the Design Expert plot below, designs 1 and 3 do not differ; however, design 2 is different than designs 1 and 3.

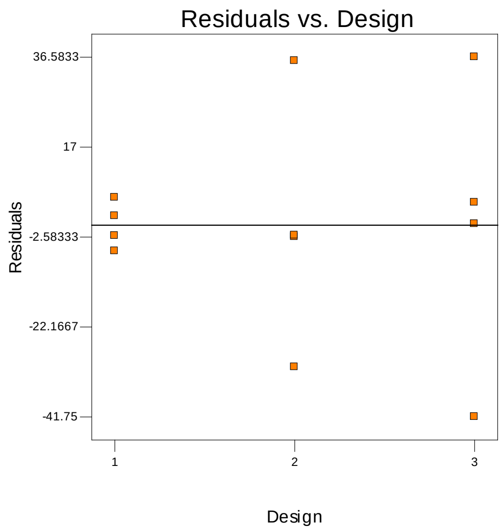


(c) Analyze the residuals from this experiment.

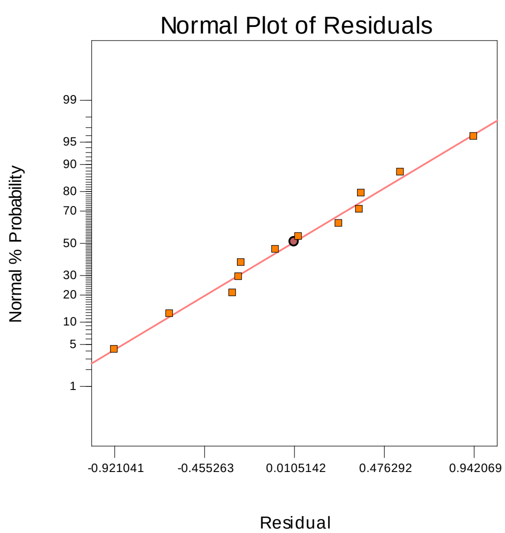
The first set of residual plots presented below represent the untransformed data. Concerns with normality as well as inequality of variance are presented. The second set of residual plots represent transformed data and do not identify significant violations of the assumptions. The residuals vs. design plot indicates a slight inequality of variance; however, not a strong violation and an improvement over the non-transformed data.

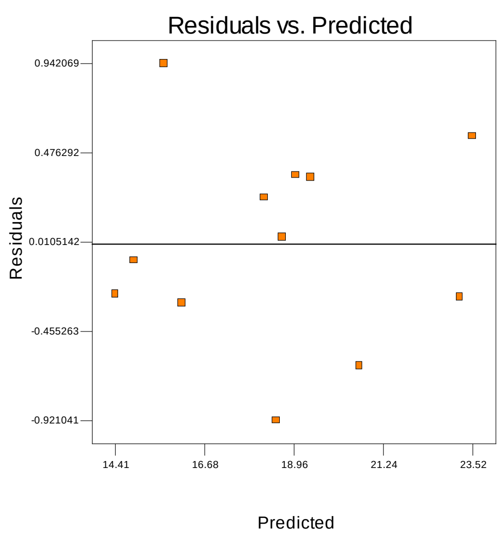


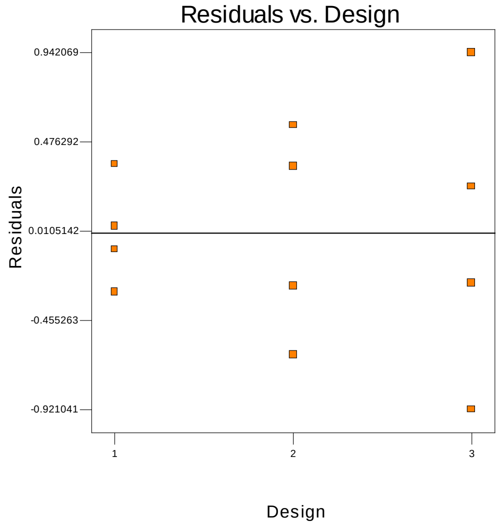




The following are the square root transformed data residual plots.







4.9. The effect of three different lubricating oils on fuel economy in diesel truck engines is being studied. Fuel economy is measured using brake-specific fuel consumption after the engine has been running for 15 minutes. Five different truck engines are available for the study, and the experimenters conduct the following randomized complete block design.

	Truck				
Oil	1	2	3	4	5
1	0.500	0.634	0.487	0.329	0.512
2	0.535	0.675	0.520	0.435	0.540
3	0.513	0.595	0.488	0.400	0.510

(a) Analyze the data from this experiment.

From the analysis below, there is a significant difference between lubricating oils with regards to fuel economy.

Design Expert Output

Response: Fuel consumption

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**	
Block	0.092	4	0.023			
Model	6.706E-003	2	3.353E-003	6.35	0.0223	significant
A	*6.706E-003*	*2*	*3.353E-003*	*6.35*	*0.0223*	
Residual	4.222E-003	8	5.278E-004			
Cor Total	0.10	14				

The Model F-value of 6.35 implies the model is significant. There is only a 2.23% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.023	R-Squared	0.6136
Mean	0.51	Adj R-Squared	0.5170
C.V.	4.49	Pred R-Squared	-0.3583
PRESS	0.015	Adeq Precision	18.814

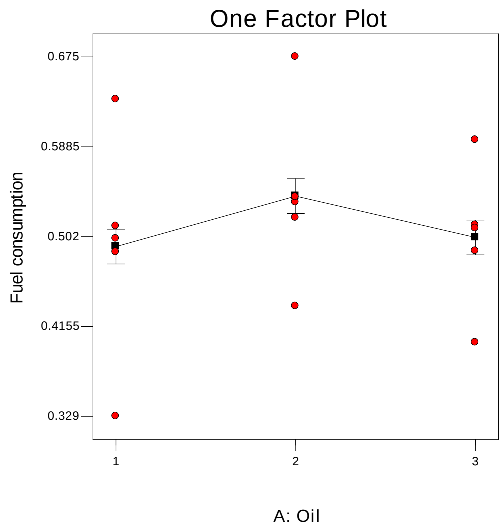
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	0.49	0.010
2-2	0.54	0.010
3-3	0.50	0.010

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-0.049	1	0.015	-3.34	0.0102
1 vs 3	-8.800E-003	1	0.015	-0.61	0.5615
2 vs 3	0.040	1	0.015	2.74	0.0255

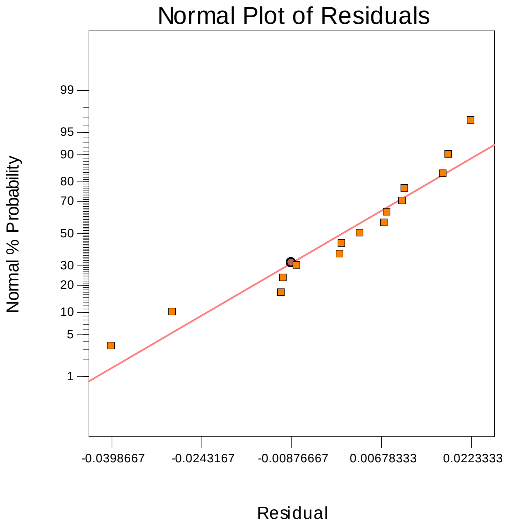
(b) Use the Fisher LSD method to make comparisons among the three lubricating oils to determine specifically which oils differ in break-specific fuel consumption.

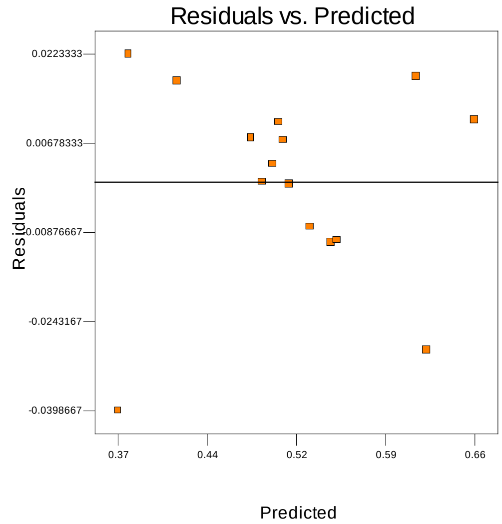
Based on the LSD bars in the Design Expert plot below, the means for break-specific fuel consumption for oils 1 and 3 do not differ; however, oil 2 is different than oils 1 and 3.

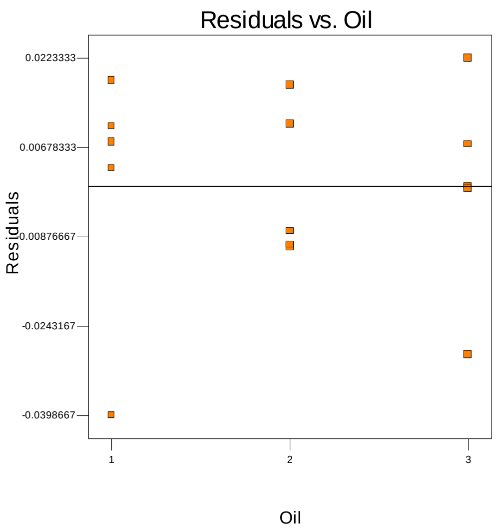


(c) Analyze the residuals from this experiment.

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







4.10. An article in the *Fire Safety Journal* ("The Effect of Nozzle Design on the Stability and Performance of Turbulent Water Jets," Vol. 4, August 1981) describes an experiment in which a shape factor was determined for several different nozzle designs at six levels of jet efflux velocity. Interest focused on potential differences between nozzle designs, with velocity considered as a nuisance variable. The data are shown below:

	Jet Efflux Velocity (m/s)					
Nozzle Design	11.73	14.37	16.59	20.43	23.46	28.74
1	0.78	0.80	0.81	0.75	0.77	0.78
2	0.85	0.85	0.92	0.86	0.81	0.83
3	0.93	0.92	0.95	0.89	0.89	0.83
4	1.14	0.97	0.98	0.88	0.86	0.83
5	0.97	0.86	0.78	0.76	0.76	0.75

(a) Does nozzle design affect the shape factor? Compare nozzles with a scatter plot and with an analysis of variance, using $\alpha = 0.05$.

Design Expert Output

Response: Shape

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	0.063	5	0.013			
Model	0.10	4	0.026	8.92	0.0003	significant
A	*0.10*	*4*	*0.026*	*8.92*	*0.0003*	
Residual	0.057	20	2.865E-003			
Cor Total	0.22	29				

The Model F-value of 8.92 implies the model is significant. There is only a 0.03% chance that a "Model F-Value" this large could occur due to noise.

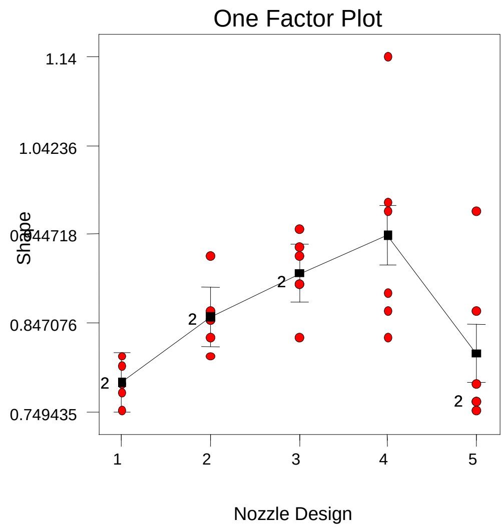
Std. Dev.	0.054	R-Squared	0.6407
Mean	0.86	Adj R-Squared	0.5688
C.V.	6.23	Pred R-Squared	0.1916
PRESS	0.13	Adeq Precision	9.438

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	0.78	0.022
2-2	0.85	0.022
3-3	0.90	0.022
4-4	0.94	0.022
5-5	0.81	0.022

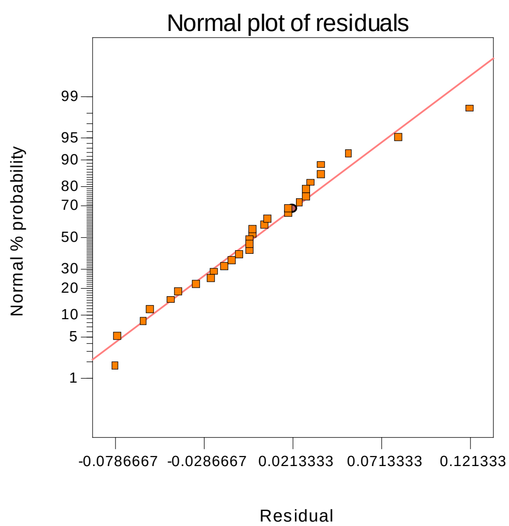
Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-0.072	1	0.031	-2.32	0.0311
1 vs 3	-0.12	1	0.031	-3.88	0.0009
1 vs 4	-0.16	1	0.031	-5.23	< 0.0001
1 vs 5	-0.032	1	0.031	-1.02	0.3177
2 vs 3	-0.048	1	0.031	-1.56	0.1335
2 vs 4	-0.090	1	0.031	-2.91	0.0086
2 vs 5	0.040	1	0.031	1.29	0.2103
3 vs 4	-0.042	1	0.031	-1.35	0.1926
3 vs 5	0.088	1	0.031	2.86	0.0097
4 vs 5	0.13	1	0.031	4.21	0.0004

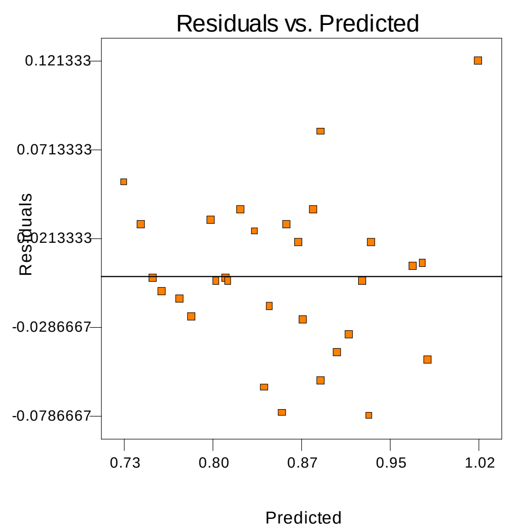
Nozzle design has a significant effect on shape factor.

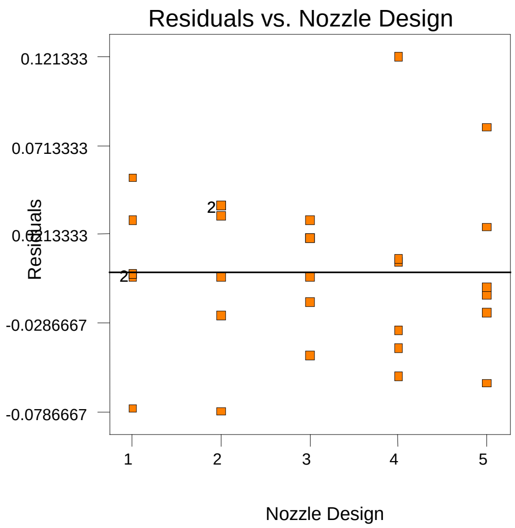


(b) Analyze the residual from this experiment.

The plots shown below do not give any indication of serious problems. Thre is some indication of a mild outlier on the normal probability plot and on the plot of residuals versus the predicted velocity.







(c) Which nozzle designs are different with respect to shape factor? Draw a graph of average shape factor for each nozzle type and compare this to a scaled *t* distribution. Compare the conclusions that you draw from this plot to those from Duncan's multiple range test.

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MS_E}{b}} = \sqrt{\frac{0.002865}{6}} = 0.021852$$

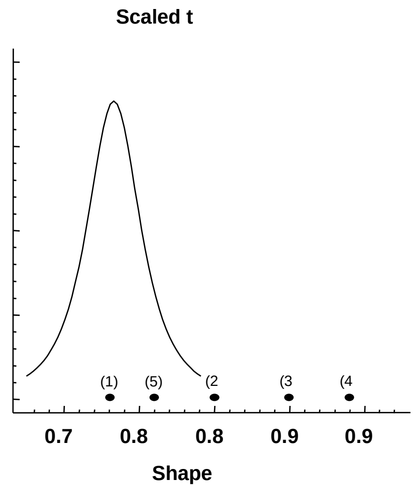
$R_2 = r_{0.05}(2,20)\ S_{\bar{y}_{i.}} = (2.95)(0.021852) = 0.06446$

$R_3 = r_{0.05}(3,20)\ S_{\bar{y}_{i.}} = (3.10)(0.021852) = 0.06774$

$R_4 = r_{0.05}(4,20)\ S_{\bar{y}_{i.}} = (3.18)(0.021852) = 0.06949$

$R_5 = r_{0.05}(5,20)\ S_{\bar{y}_{i.}} = (3.25)(0.021852) = 0.07102$

	Mean Difference		*R*	
1 vs 4	0.16167	>	0.07102	different
1 vs 3	0.12000	>	0.06949	different
1 vs 2	0.07167	>	0.06774	different
1 vs 5	0.03167	<	0.06446	
5 vs 4	0.13000	>	0.06949	different
5 vs 3	0.08833	>	0.06774	different
5 vs 2	0.04000	<	0.06446	
2 vs 4	0.09000	>	0.06774	different
2 vs 3	0.04833	<	0.06446	
3 vs 4	0.04167	<	0.06446	



4.11. An article in *Communications of the ACM* (Vol. 30, No. 5, 1987) studied different algorithms for estimating software development costs. Six algorithms were applied to several different software development projects and the percent error in estimating the development cost was observed. Some of the data from this experiment is show in the table below.

	Project					
Algorithm	1	2	3	4	5	6
1 (SLIM)	1244	21	82	2221	905	839
2 (COCOMO-A)	281	129	396	1306	336	910
3 (COCOMO-R)	220	84	458	543	300	794
4 (COCOMO-C)	225	83	425	552	291	826
5 (FUNCTION POINTS)	19	11	-34	121	15	103
6 (ESTIMALS)	-20	35	-53	170	104	199

(a) Do the algorithms differ in their mean cost estimation accuracy?

The ANOVA below identifies the algorithms are significantly different in their mean cost estimation error.

Design Expert Output

Response **Cost Error**
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Block	2.287E+006	5	4.575E+005			
Model	2.989E+006	5	5.978E+005	5.38	0.0017	significant
A-Algorithm	*2.989E+006*	*5*	*5.978E+005*	*5.38*	*0.0017*	
Residual	2.780E+006	25	1.112E+005			
Cor Total	8.056E+006	35				

The Model F-value of 5.38 implies the model is significant. There is only a 0.17% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	333.44	R-Squared	0.5182
Mean	392.81	Adj R-Squared	0.4218
C.V. %	84.89	Pred R-Squared	0.0009
PRESS	5.764E+006	Adeq Precision	8.705

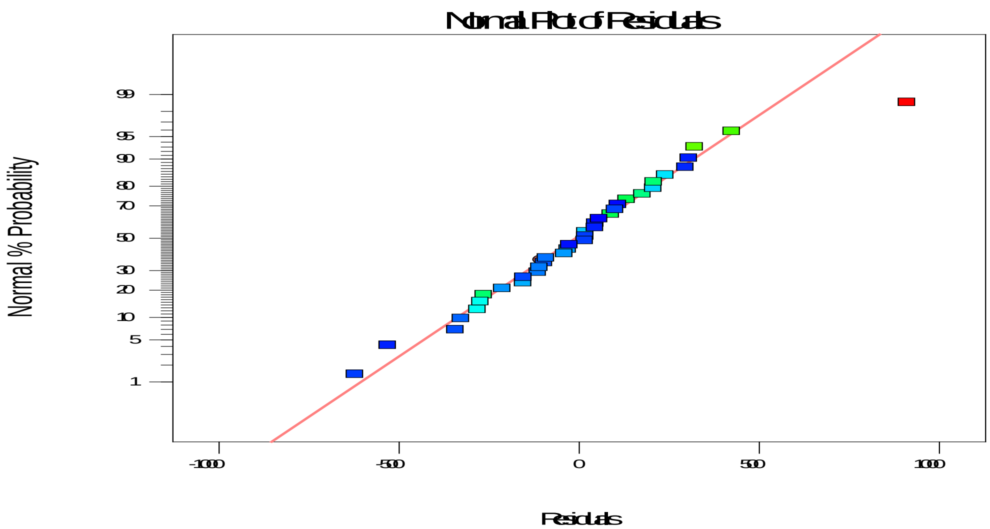
Treatment Means (Adjusted, If Necessary)

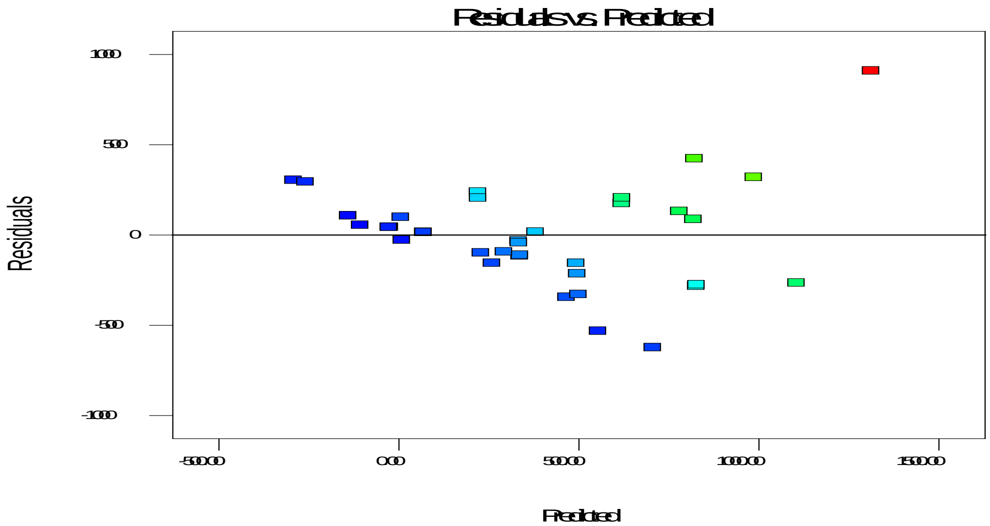
	Estimated Mean	Standard Error
1-SLM	885.33	136.13
2-COCOMO-A	559.67	136.13
3-COCOMO-R	399.83	136.13
4-COCOMO-C	400.33	136.13
5-FUNCTION POINTS	39.17	136.13
6-ESTIMALS	72.50	136.13

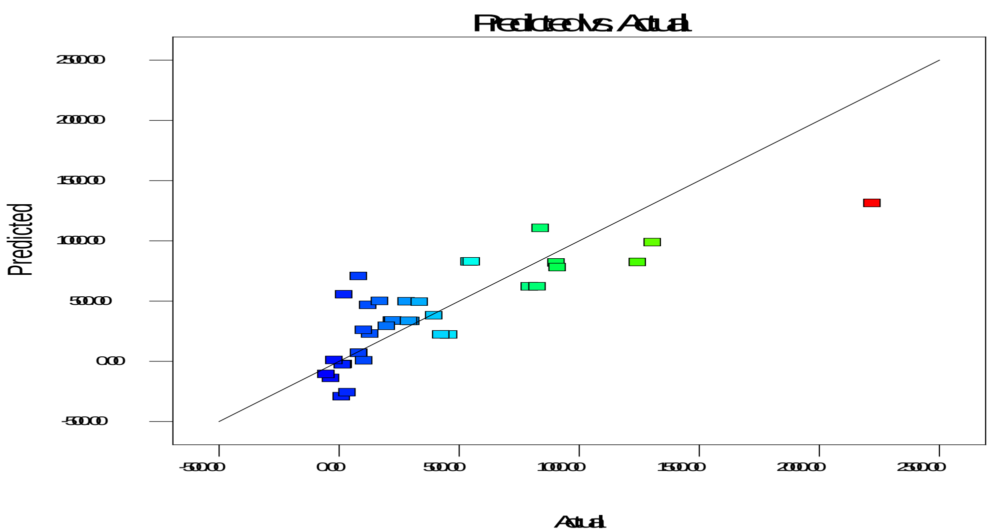
Treatment	Mean Difference	df	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	325.67	1	192.51	1.69	0.1031
1 vs 3	485.50	1	192.51	2.52	0.0184
1 vs 4	485.00	1	192.51	2.52	0.0185
1 vs 5	846.17	1	192.51	4.40	0.0002
1 vs 6	812.83	1	192.51	4.22	0.0003
2 vs 3	159.83	1	192.51	0.83	0.4143
2 vs 4	159.33	1	192.51	0.83	0.4157
2 vs 5	520.50	1	192.51	2.70	0.0122
2 vs 6	487.17	1	192.51	2.53	0.0181
3 vs 4	-0.50	1	192.51	-2.597E-003	0.9979
3 vs 5	360.67	1	192.51	1.87	0.0727
3 vs 6	327.33	1	192.51	1.70	0.1015
4 vs 5	361.17	1	192.51	1.88	0.0724
4 vs 6	327.83	1	192.51	1.70	0.1010
5 vs 6	-33.33	1	192.51	-0.17	0.8639

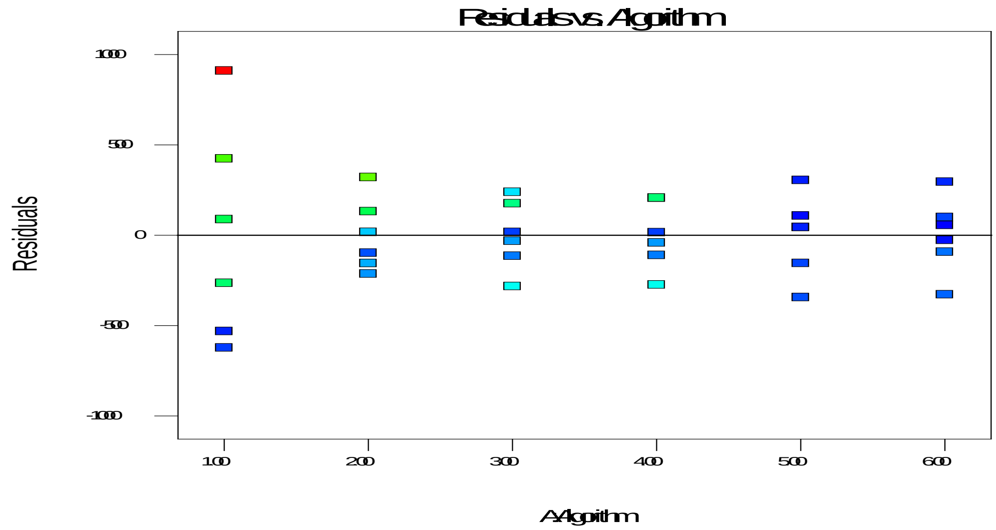
(b) Analyze the residuals from this experiment.

The residual plots below identify a single outlier that should be investigated.




Residuals vs. Predicted
Residuals
Predicted


Predicted vs. Actual
Predicted
Actual



(c) Which algorithm would you recommend for use in practice?

The FUNCTIONAL POINTS algorithm has the losest cost estimation error.

4.12. An article in *Nature Genetics* (2003, Vol. 34, pp. 85-90) "Treatment-Specific Changes in Gene Expression Discriminate in vivo Drug Response in Human Leukemia Cells" studied gene expressionas a function of different treatments for leukemia. Three treatment groups are: mercaptopurine (MP) only; low-dose methotrexate (LDMTX) and MP; and high-dose methotrexate (HDMTX) and MP. Each group contained ten subjects. The responses from a specific gene are shown in the table below:

	Project									
MP ONLY	334.5	31.6	701	41.2	61.2	69.6	67.5	66.6	120.7	881.9
MP + HDMTX	919.4	404.2	1024.8	54.1	62.8	671.6	882.1	354.2	321.9	91.1
MP + LDMTX	108.4	26.1	240.8	191.1	69.7	242.8	62.7	396.9	23.6	290.4

(a) Is there evidence to support the claim that the treatment means differ?

The ANOVA below identifies the treatment means are significantly different.

Design Expert Output

Response **Gene Expression**
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Block	9.206E+005	9	1.023E+005			
Model	5.384E+005	2	2.692E+005	3.68	0.0457	significant
A-Treatment	*5.384E+005*	*2*	*2.692E+005*	*3.68*	*0.0457*	
Residual	1.316E+006	18	73130.15			
Cor Total	2.775E+006	29				

The Model F-value of 3.68 implies the model is significant. There is only a 4.57% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	270.43	R-Squared	0.2903
Mean	293.82	Adj R-Squared	0.2114
C.V. %	92.04	Pred R-Squared	-0.9714
PRESS	3.657E+006	Adeq Precision	5.288

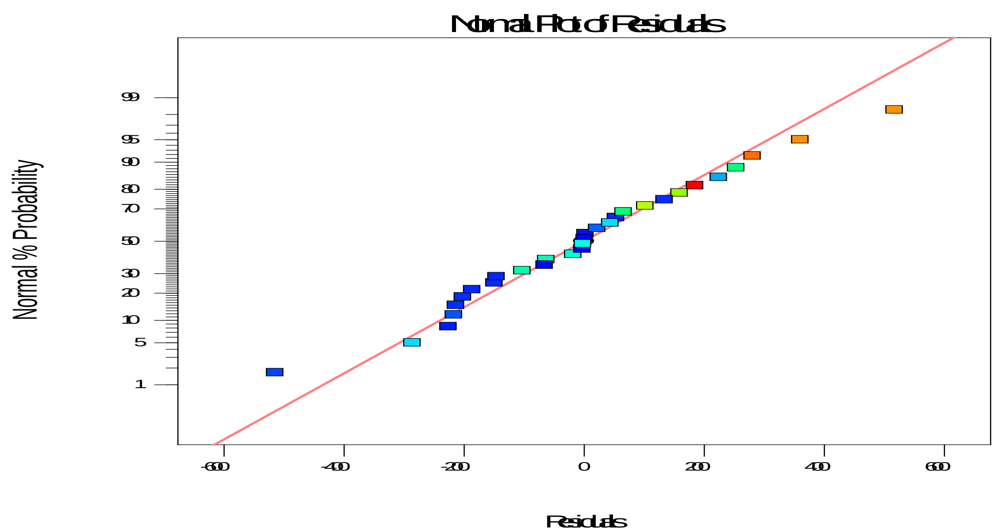
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-MP Only	237.58	85.52
2-MP + HDMTX	478.62	85.52
3-MP + LDMTX	165.25	85.52

Treatment	Mean Difference	df	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	-241.04	1	120.94	-1.99	0.0616
1 vs 3	72.33	1	120.94	0.60	0.5572
2 vs 3	313.37	1	120.94	2.59	0.0184

(b) Chec the normality assumption. Can we assume these samples are from normal populations?

The normal plot of residuals below identifies a slightly non-normal distribution.



(c) Take the logarithm of the raw data. Is there evidence to support the claim that the treatment means differ for the transformed data?

The ANOVA for the natural log transformed data identifies the treatment means as only moderately different with an *F* value of 0.07

Design Expert Output

Response **Gene Expression**
Transform:Natural Log Constant: **0**
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Block	14.75	9	1.64		
Model	6.30	2	3.15	3.09	0.0700
A-Treatment	*6.30*	*2*	*3.15*	*3.09*	*0.0700*
Residual	18.32	18	1.02		
Cor Total	39.37	29			

The Model F-value of 3.09 implies there is a 7.00% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.01	R-Squared	0.2558
Mean	5.09	Adj R-Squared	0.1731
C.V. %	19.83	Pred R-Squared	-1.0672
PRESS	50.89	Adeq Precision	4.942

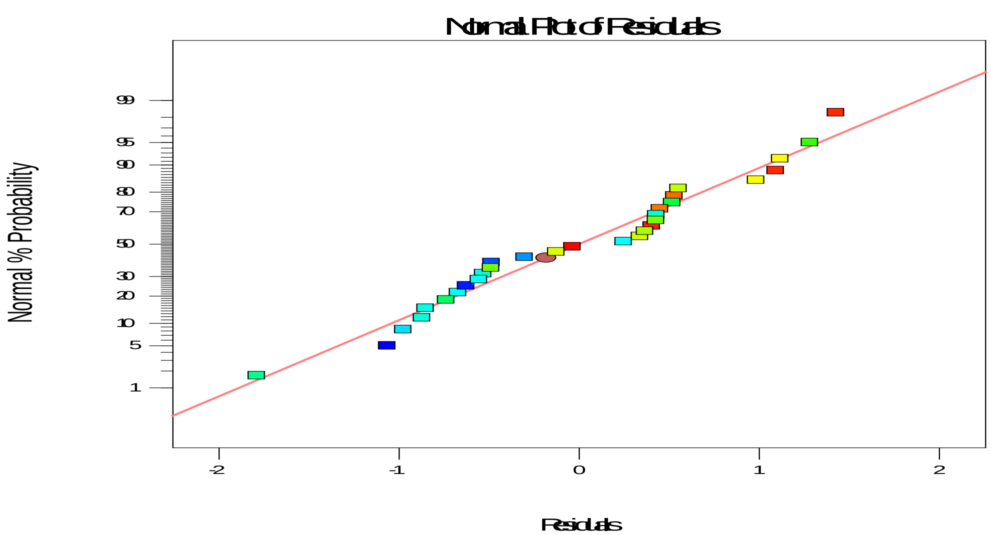
Treatment Means (Adjusted, If Necessary)

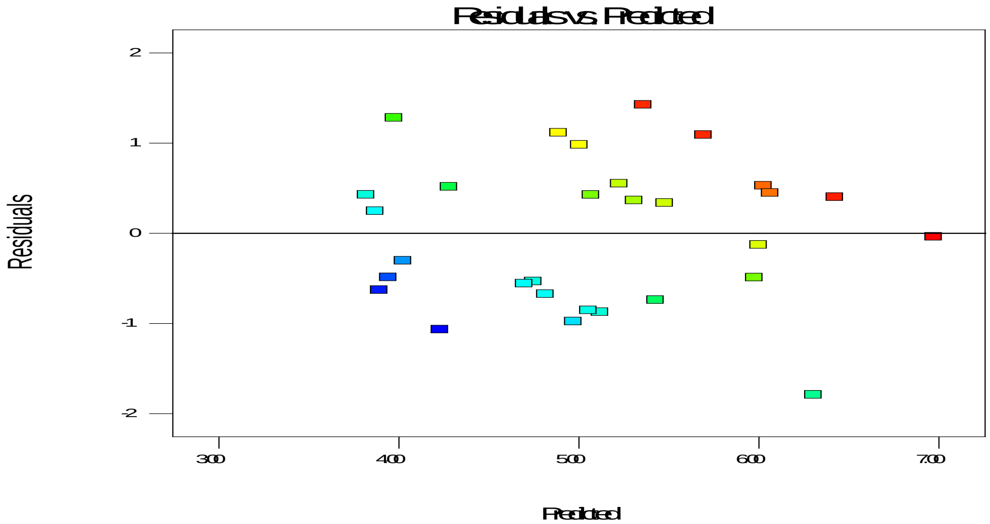
	Estimated Mean	**Standard Error**
1-MP Only	4.79	0.32
2-MP + HDMTX	5.73	0.32
3-MP + LDMTX	4.74	0.32

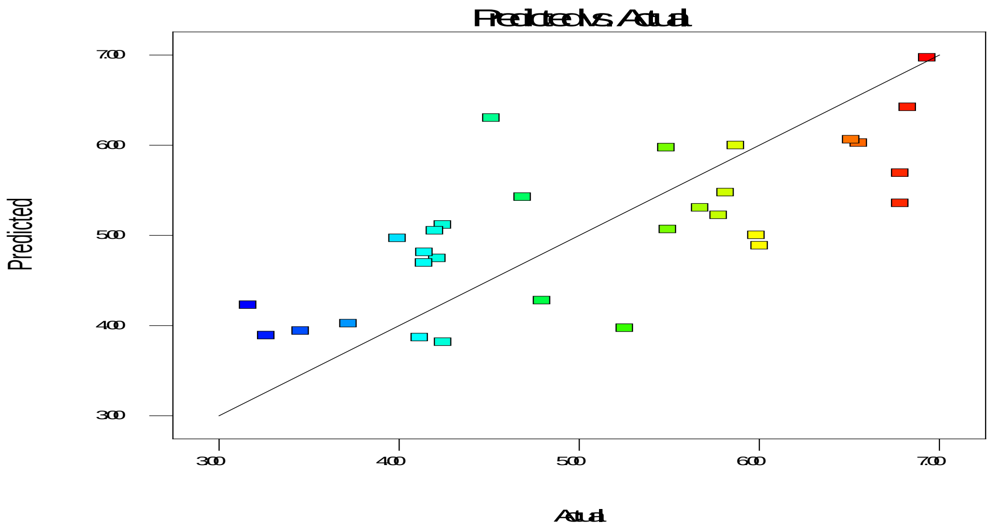
Treatment	**Mean Difference**	**df**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-0.95	1	0.45	-2.10	0.0505
1 vs 3	0.050	1	0.45	0.11	0.9122
2 vs 3	1.00	1	0.45	2.21	0.0405

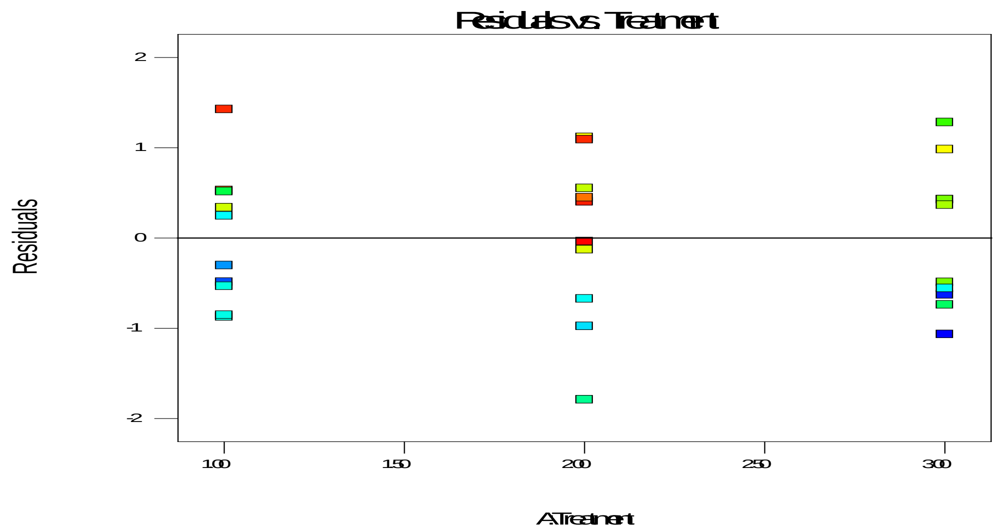
(d) Analyze the residuals from the transformed data and comment on model adequacy.

The residual plots below identify no concerns with the model adequacy.




Residuals vs. Predicted
Residuals
2
1
0
-1
-2
300
400
500
600
700
Predicted


Predicted vs. Actual
Predicted
700
600
500
400
300
300
400
500
600
700
Actual



4.13. Consider the ratio control algorithm experiment described in Section 3.8. The experiment was actually conducted as a randomized block design, where six time periods were selected as the blocks, and all four ratio control algorithms were tested in each time period. The average cell voltage and the standard deviation of voltage (shown in parentheses) for each cell are as follows:

Ratio Control	Time Period					
Algorithms	1	2	3	4	5	6
1	4.93 (0.05)	4.86 (0.04)	4.75 (0.05)	4.95 (0.06)	4.79 (0.03)	4.88 (0.05)
2	4.85 (0.04)	4.91 (0.02)	4.79 (0.03)	4.85 (0.05)	4.75 (0.03)	4.85 (0.02)
3	4.83 (0.09)	4.88 (0.13)	4.90 (0.11)	4.75 (0.15)	4.82 (0.08)	4.90 (0.12)
4	4.89 (0.03)	4.77 (0.04)	4.94 (0.05)	4.86 (0.05)	4.79 (0.03)	4.76 (0.02)

(a) Analyze the average cell voltage data. (Use α = 0.05.) Does the choice of ratio control algorithm affect the cell voltage?

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**	
Block	0.017	5	3.487E-003			
Model	2.746E-003	3	9.153E-004	0.19	0.9014	not significant
A	*2.746E-003*	*3*	*9.153E-004*	*0.19*	*0.9014*	
Residual	0.072	15	4.812E-003			
Cor Total	0.092	23				

The "Model F-value" of 0.19 implies the model is not significant relative to the noise. There is a 90.14 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	0.069	R-Squared	0.0366
Mean	4.84	Adj R-Squared	-0.1560
C.V.	1.43	Pred R-Squared	-1.4662
PRESS	0.18	Adeq Precision	2.688

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**

1-1	4.86	0.028
2-2	4.83	0.028
3-3	4.85	0.028
4-4	4.84	0.028

Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	0.027	1	0.040	0.67	0.5156
1 vs 3	0.013	1	0.040	0.33	0.7438
1 vs 4	0.025	1	0.040	0.62	0.5419
2 vs 3	-0.013	1	0.040	-0.33	0.7438
2 vs 4	-1.667E-003	1	0.040	-0.042	0.9674
3 vs 4	0.012	1	0.040	0.29	0.7748

The ratio control algorithm does not affect the mean cell voltage.

(b) Perform an appropriate analysis of the standard deviation of voltage. (Recall that this is called "pot noise.") Does the choice of ratio control algorithm affect the pot noise?

Design Expert Output

Response: StDev 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	
Block	0.94	5	0.19			
Model	6.17	3	2.06	33.26	< 0.0001	significant
A	*6.17*	*3*	*2.06*	*33.26*	*< 0.0001*	
Residual	0.93	15	0.062			
Cor Total	8.04	23				

The Model F-value of 33.26 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.25	R-Squared	0.8693
Mean	-3.04	Adj R-Squared	0.8432
C.V.	-8.18	Pred R-Squared	0.6654
PRESS	2.37	Adeq Precision	12.446

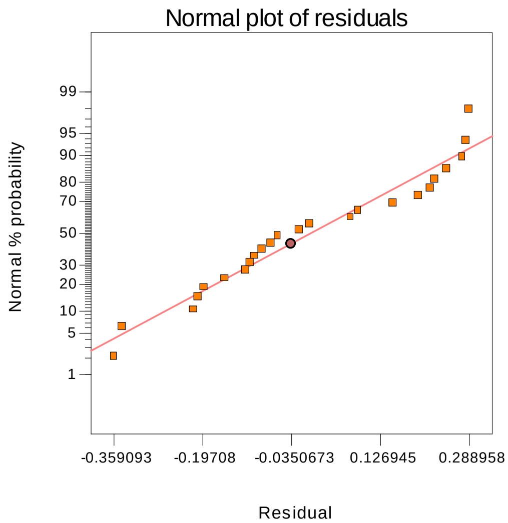
Treatment Means (Adjusted, If Necessary)

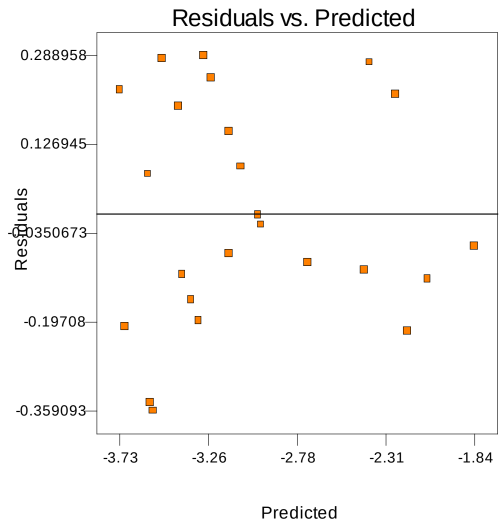
	Estimated Mean	Standard Error
1-1	-3.09	0.10
2-2	-3.51	0.10
3-3	-2.20	0.10
4-4	-3.36	0.10

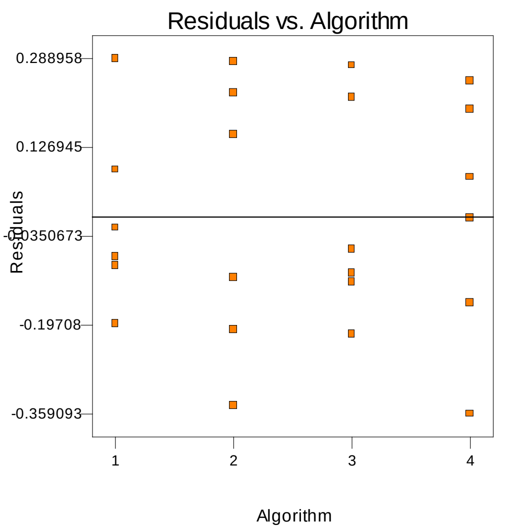
Treatment	Mean Difference	DF	Standard Error	t for H0 Coeff=0	Prob > |t|
1 vs 2	0.42	1	0.14	2.93	0.0103
1 vs 3	-0.89	1	0.14	-6.19	< 0.0001
1 vs 4	0.27	1	0.14	1.87	0.0813
2 vs 3	-1.31	1	0.14	-9.12	< 0.0001
2 vs 4	-0.15	1	0.14	-1.06	0.3042
3 vs 4	1.16	1	0.14	8.06	< 0.0001

A natural log transformation was applied to the pot noise data. The ratio control algorithm does affect the pot noise.

(c) Conduct any residual analyses that seem appropriate.







The normal probability plot shows slight deviations from normality; however, still acceptable.

(d) Which ratio control algorithm would you select if your objective is to reduce both the average cell voltage and the pot noise?

Since the ratio control algorithm has little effect on average cell voltage, select the algorithm that minimizes pot noise, that is algorithm #2.

4.14. An aluminum master alloy manufacturer produces grain refiners in ingot form. The company produces the product in four furnaces. Each furnace is known to have its own unique operating characteristics, so any experiment run in the foundry that involves more than one furnace will consider furnaces as a nuisance variable. The process engineers suspect that stirring rate impacts the grain size of the product. Each furnace can be run at four different stirring rates. A randomized block design is run for a particular refiner and the resulting grain size data is as follows.

	Furnace			
Stirring Rate	1	2	3	4
5	8	4	5	6
10	14	5	6	9
15	14	6	9	2
20	17	9	3	6

(a) Is there any evidence that stirring rate impacts grain size?

Design Expert Output

Response: Grain Size

ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	165.19	3	55.06			
Model	22.19	3	7.40	0.85	0.4995	not significant
A	*22.19*	*3*	*7.40*	*0.85*	*0.4995*	
Residual	78.06	9	8.67			
Cor Total	265.44	15				

The "Model F-value" of 0.85 implies the model is not significant relative to the noise. There is a 49.95 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	2.95	R-Squared	0.2213
Mean	7.69	Adj R-Squared	-0.0382
C.V.	38.31	Pred R-Squared	-1.4610
PRESS	246.72	Adeq Precision	5.390

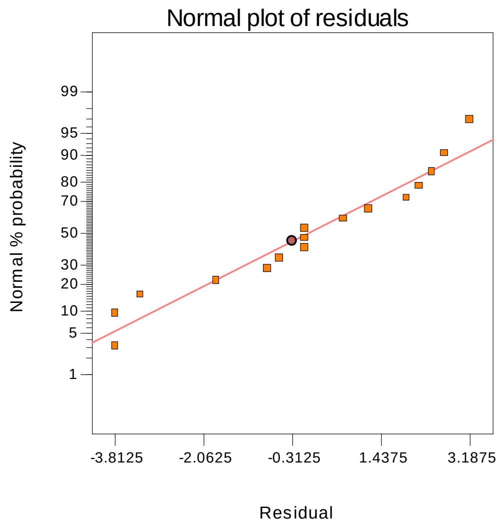
Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-5	5.75	1.47
2-10	8.50	1.47
3-15	7.75	1.47
4-20	8.75	1.47

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-2.75	1	2.08	-1.32	0.2193
1 vs 3	-2.00	1	2.08	-0.96	0.3620
1 vs 4	-3.00	1	2.08	-1.44	0.1836
2 vs 3	0.75	1	2.08	0.36	0.7270
2 vs 4	-0.25	1	2.08	-0.12	0.9071
3 vs 4	-1.00	1	2.08	-0.48	0.6425

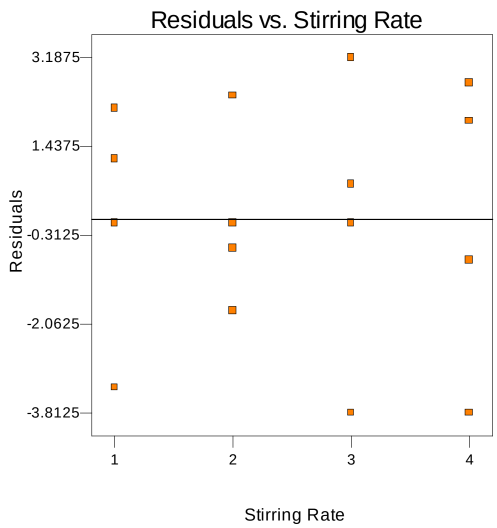
The analysis of variance shown above indicates that there is no difference in mean grain size due to the different stirring rates.

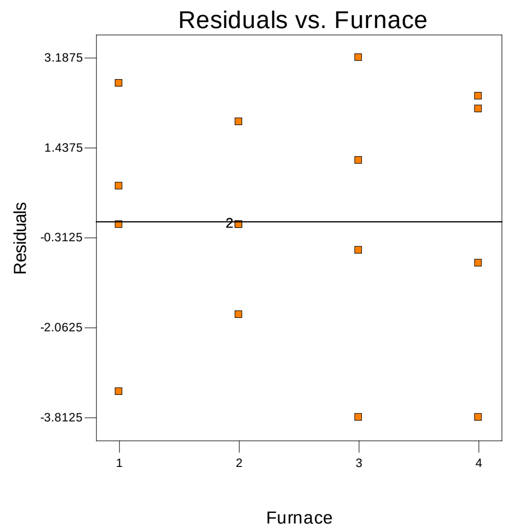
(b) Graph the residuals from this experiment on a normal probability plot. Interpret this plot.



The plot indicates that normality assumption is valid.

(c) Plot the residuals versus furnace and stirring rate. Does this plot convey any useful information?





The variance is consistent at different stirring rates. Not only does this validate the assumption of uniform variance, it also identifies that the different stirring rates do not affect variance.

(d) What should the process engineers recommend concerning the choice of stirring rate and furnace for this particular grain refiner if small grain size is desirable?

There really is no effect due to the stirring rate.

4.15. Analyze the data in Problem 4.4 using the general regression significance test.

$$
\begin{array}{llllllllll}
\mu: & 12\hat{\mu} & +4\hat{\tau}_1 & +4\hat{\tau}_2 & +4\hat{\tau}_3 & +3\hat{\beta}_1 & +3\hat{\beta}_2 & +3\hat{\beta}_3 & +3\hat{\beta}_4 & =225 \\
\tau_1: & 4\hat{\mu} & +4\hat{\tau}_1 & & & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =92 \\
\tau_2: & 4\hat{\mu} & & +4\hat{\tau}_2 & & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =101 \\
\tau_3: & 4\hat{\mu} & & & +4\hat{\tau}_3 & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =32 \\
\beta_1: & 3\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & +3\hat{\beta}_1 & & & & =34 \\
\beta_1: & 3\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & & +3\hat{\beta}_2 & & & =50 \\
\beta_1: & 3\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & & & +3\hat{\beta}_3 & & =36 \\
\beta\ : & 3\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & & & & +3\hat{\beta}_4 & =105
\end{array}
$$

Applying the constraints $\sum \hat{\tau}_i = \sum \hat{\beta}_j = 0$, we obtain:

$$\hat{\mu} = \frac{225}{12},\ \hat{\tau}_1 = \frac{51}{12},\ \hat{\tau}_2 = \frac{78}{12},\ \hat{\tau}_3 = \frac{-129}{12},\ \hat{\beta}_1 = \frac{-89}{12},\ \hat{\beta}_2 = \frac{-25}{12},\ \hat{\beta}_3 = \frac{-81}{12},\ \hat{\beta}_4 = \frac{195}{12}$$

$$R(\mu,\tau,\beta) = \left(\frac{225}{12}\right)(225) + \left(\frac{51}{12}\right)(92) + \left(\frac{78}{12}\right)(101) + \left(\frac{-129}{12}\right)(32) + \left(\frac{-89}{12}\right)(34) + \left(\frac{-25}{12}\right)(50) + \left(\frac{-81}{12}\right)(36) + \left(\frac{195}{12}\right)(105) = 6029.17$$

$$\sum\sum y_{ij}^2 = 6081,\ SS_E = \sum\sum y_{ij}^2 - R(\mu,\tau,\beta) = 6081 - 6029.17 = 51.83$$

Model Restricted to $\tau_i = 0$:

$$
\begin{array}{lllllll}
\mu: & 12\hat{\mu} & +3\hat{\beta}_1 & +3\hat{\beta}_2 & +3\hat{\beta}_3 & +3\hat{\beta}_4 & =225 \\
\beta_1: & 3\hat{\mu} & +3\hat{\beta}_1 & & & & =34 \\
\beta_2: & 3\hat{\mu} & & +3\hat{\beta}_2 & & & =50 \\
\beta_3: & 3\hat{\mu} & & & +3\hat{\beta}_3 & & =36 \\
\beta_4: & 3\hat{\mu} & & & & +3\hat{\beta}_4 & =105
\end{array}
$$

Applying the constraint $\sum \hat{\beta}_j = 0$, we obtain:

$$\hat{\mu} = \frac{225}{12},\ \hat{\beta}_1 = -89/12,\ \hat{\beta}_2 = \frac{-25}{12},\ \hat{\beta}_3 = \frac{-81}{12},\ \hat{\beta}_4 = \frac{195}{12}.\ \text{Now:}$$

$$R(\mu,\beta) = \left(\frac{225}{12}\right)(225) + \left(\frac{-89}{12}\right)(34) + \left(\frac{-25}{12}\right)(50) + \left(\frac{-81}{12}\right)(36) + \left(\frac{195}{12}\right)(105) = 5325.67$$

$$R(\tau|\mu,\beta) = R(\mu,\tau,\beta) - R(\mu,\beta) = 6029.17 - 5325.67 = 703.50 = SS_{Treatments}$$

Model Restricted to $\beta_j = 0$:

$$
\begin{array}{llllll}
\mu: & 12\hat{\mu} & +4\hat{\tau}_1 & +4\hat{\tau}_2 & +4\hat{\tau}_3 & =225 \\
\tau_1: & 4\hat{\mu} & +4\hat{\tau}_1 & & & =92 \\
\tau_2: & 4\hat{\mu} & & +4\hat{\tau}_2 & & =101 \\
\tau_3: & 4\hat{\mu} & & & +4\hat{\tau}_3 & =32
\end{array}
$$

Applying the constraint $\sum \hat{\tau}_i = 0$, we obtain:

$$\hat{\mu} = \frac{225}{12},\ \hat{\tau}_1 = \frac{51}{12},\ \hat{\tau}_2 = \frac{78}{12},\ \hat{\tau}_3 = \frac{-129}{12}$$

$$R(\mu,\tau) = \left(\frac{225}{12}\right)(225) + \left(\frac{51}{12}\right)(92) + \left(\frac{78}{12}\right)(101) + \left(\frac{-129}{12}\right)(32) = 4922.25$$

$$R(\beta|\mu,\tau) = R(\mu,\tau,\beta) - R(\mu,\tau) = 6029.17 - 4922.25 = 1106.92 = SS_{Blocks}$$

4.16. Assuming that chemical types and bolts are fixed, estimate the model parameters τ_i and β_j in Problem 4.3.

Using Equations 4.18, applying the constraints, we obtain:

$$\hat{\mu} = \frac{35}{20},\ \hat{\tau}_1 = \frac{-23}{20},\ \hat{\tau}_2 = \frac{-7}{20},\ \hat{\tau}_3 = \frac{13}{20},\ \hat{\tau}_4 = \frac{17}{20},\ \hat{\beta}_1 = \frac{35}{20},\ \hat{\beta}_2 = \frac{-65}{20},\ \hat{\beta}_3 = \frac{75}{20},\ \hat{\beta}_4 = \frac{20}{20},\ \hat{\beta}_5 = \frac{-65}{20}$$

4.17. Draw an operating characteristic curve for the design in Problem 4.4. Does this test seem to be sensitive to small differences in treatment effects?

Assuming that solution type is a fixed factor, we use the OC curve in appendix V. Calculate

$$\Phi^2 = \frac{b\sum \tau_i^2}{a\sigma^2} = \frac{4\sum \tau_i^2}{3(8.64)}$$

using MS_E to estimate σ^2. We have:

$$\upsilon_1 = a - 1 = 2 \qquad \upsilon_2 = (a-1)(b-1) = (2)(3) = 6\ .$$

If $\sum \hat{\tau}_i^2 = \sigma^2 = MS_E$, then:

$$\Phi = \sqrt{\frac{4}{3(1)}} = 1.15 \text{ and } \beta \cong 0.70$$

If $\sum \hat{\tau}_i = 2\sigma^2 = 2MS_E$, then:

$$\Phi = \sqrt{\frac{4(2)}{3(1)}} = 1.63 \text{ and } \beta \cong 0.55\text{, etc.}$$

This test is not very sensitive to small differences.

4.18. Suppose that the observation for chemical type 2 and bolt 3 is missing in Problem 4.3. Analyze the problem by estimating the missing value. Perform the exact analysis and compare the results.

$$y_{23} \text{ is missing.} \quad \hat{y}_{23} = \frac{ay_{2.}^{'} + by_{.3}^{'} - y_{..}^{'}}{(a-1)(b-1)} = \frac{4(282)+5(227)-1360}{(3)(4)} = 75.25$$

Therefore, $y_{2.}$=357.25, $y_{.3}$=302.25, and $y_{..}$=1435.25

Source	*SS*	*DF*	*MS*	F_0
Chemicals	12.7844	3	4.2615	2.154
Bolts	158.8875	4		
Error	21.7625	11	1.9784	
Total	193.4344	18		

$F_{0.05,3,11}$=3.59, Chemicals are not significant. This is the same result as found in Problem 4.3.

4.19. Consider the hardness testing experiment in Problem 4.7. Suppose that the observation for tip 2 in coupon 3 is missing. Analyze the problem by estimating the missing value.

$$y_{23} \text{ is missing.} \quad \hat{y}_{23} = \frac{ay_{2.}^{'} + by_{.3}^{'} - y_{..}^{'}}{(a-1)(b-1)} = \frac{4(28.6)+4(29.1)-144.2}{(3)(3)} = 9.62$$

Therefore, $y_{2.}$=38.22, $y_{.3}$=38.72, and $y_{..}$=153.82

Source	*SS*	*DF*	*MS*	F_0
Tip	0.40	3	0.133333	19.29
Coupon	0.80	3		
Error	0.0622	9	0.006914	
Total	1.2622	15		

$F_{0.05,3,9}$=3.86, Tips are significant. This is the same result as found in Problem 4.7.

4.20. ***Two missing values in a randomized block.*** Suppose that in Problem 4.3 the observations for chemical type 2 and bolt 3 and chemical type 4 and bolt 4 are missing.

(a) Analyze the design by iteratively estimating the missing values as described in Section 4.1.3.

$$\hat{y}_{23} = \frac{4y_{2.}^{'} + 5y_{.3}^{'} - y_{..}^{'}}{12} \text{ and } \hat{y}_{44} = \frac{4y_{4.}^{'} + 5y_{.4}^{'} - y_{..}^{'}}{12}$$

Data is coded *y*-70. As an initial guess, set y_{23}^{0} equal to the average of the observations available for chemical 2. Thus, $y_{23}^{0} = \frac{2}{4} = 0.5$. Then ,

$$\hat{y}_{44}^{0} = \frac{4(8)+5(6)-25.5}{12} = 3.04$$

$$\hat{y}_{23}^{1} = \frac{4(2)+5(17)-28.04}{12} = 5.41$$

$$\hat{y}_{44}^{1} = \frac{4(8)+5(6)-30.41}{12} = 2.63$$
$$\hat{y}_{44}^{2} = \frac{4(2)+5(17)-27.63}{12} = 5.44$$
$$\hat{y}_{44}^{2} = \frac{4(8)+5(6)-30.44}{12} = 2.63$$
$$\therefore \hat{y}_{23} = 5.44 \quad \hat{y}_{44} = 2.63$$

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**	
Block	156.83	4	39.21			
Model	9.59	3	3.20	2.08	0.1560	not significant
A	*9.59*	*3*	*3.20*	*2.08*	*0.1560*	
Residual	18.41	12	1.53			
Cor Total	184.83	19				

(b) Differentiate SS_E with respect to the two missing values, equate the results to zero, and solve for estimates of the missing values. Analyze the design using these two estimates of the missing values.

$$SS_E = \sum\sum y_{ij}^2 - \tfrac{1}{5}\sum y_{i.}^2 - \tfrac{1}{4}\sum y_{.j}^2 + \tfrac{1}{20}\sum y_{..}^2$$
$$SS_E = 0.6y_{23}^2 + 0.6y_{44}^2 - 6.8y_{23} - 3.7y_{44} + 0.1y_{23}y_{44} + R$$

From $\frac{\partial SS_E}{\partial y_{23}} = \frac{\partial SS_E}{\partial y_{44}} = 0$, we obtain:

$$\begin{aligned} 1.2\hat{y}_{23} + 0.1\hat{y}_{44} &= 6.8 \\ 0.1\hat{y}_{23} + 1.2\hat{y}_{44} &= 3.7 \end{aligned} \quad \Rightarrow \hat{y}_{23} = 5.45,\ \hat{y}_{44} = 2.63$$

These quantities are almost identical to those found in part (a). The analysis of variance using these new data does not differ substantially from part (a).

(c) Derive general formulas for estimating two missing values when the observations are in *different* blocks.

$$SS_E = y_{iu}^2 + y_{kv}^2 - \frac{\left(y'_{i.} + y_{iu}\right)^2 + \left(y'_{k.} + y_{kv}\right)^2}{b} - \frac{\left(y'_{.u} + y_{iu}\right)^2 + \left(y'_{.v} + y_{kv}\right)^2}{a} + \frac{\left(y'_{..} + y_{iu} + y_{kv}\right)^2}{ab}$$

From $\frac{\partial SS_E}{\partial y_{23}} = \frac{\partial SS_E}{\partial y_{44}} = 0$, we obtain:

$$\hat{y}_{iu}\left[\frac{(a-1)(b-1)}{ab}\right] = \frac{ay'_{i.} + by'_{.j} - y'_{..}}{ab} - \frac{\hat{y}_{kv}}{ab}$$
$$\hat{y}_{kv}\left[\frac{(a-1)(b-1)}{ab}\right] = \frac{ay'_{k.} + by'_{.v} - y'_{..}}{ab} - \frac{\hat{y}_{iu}}{ab}$$

whose simultaneous solution is:

$$\hat{y}_{iu} = \frac{y'_{i.}\, a\left[1-(a-1)^2(b-1)^2-ab\right]+y'_{.u}\, b\left[1-(a-1)^2(b-1)^2-ab\right]-y'_{..}\left[1-ab(a-1)^2(b-1)^2\right]}{(a-1)(b-1)\left[1-(a-1)^2(b-1)^2\right]}+$$

$$\frac{ab\left[ay'_{k.}+by'_{.v}-y'_{..}\right]}{\left[1-(a-1)^2(b-1)^2\right]}$$

$$\hat{y}_{kv} = \frac{ay'_{i.}+by'_{.u}-y'_{..}-(b-1)(a-1)\left[ay'_{k.}+by'_{.v}-y'_{..}\right]}{\left[1-(a-1)^2(b-1)^2\right]}$$

(d) Derive general formulas for estimating two missing values when the observations are in the *same* block. Suppose that two observations y_{ij} and y_{kj} are missing, $i \neq k$ (same block j).

$$SS_E = y_{ij}^2 + y_{kj}^2 - \frac{\left(y'_{i.}+y_{ij}\right)^2+\left(y'_{k.}+y_{kj}\right)^2}{b} - \frac{\left(y'_{.j}+y_{ij}+y_{kj}\right)^2}{a} + \frac{\left(y'_{..}+y_{ij}+y_{kj}\right)^2}{ab}$$

From $\frac{\partial SS_E}{\partial y_{ij}} = \frac{\partial SS_E}{\partial y_{kj}} = 0$, we obtain

$$\hat{y}_{ij} = \frac{ay'_{i.}+by'_{.j}-y'_{..}}{(a-1)(b-1)} + \hat{y}_{kj}(a-1)(b-1)^2$$

$$\hat{y}_{kj} = \frac{ay'_{k.}+by'_{.j}-y'_{..}}{(a-1)(b-1)} + \hat{y}_{ij}(a-1)(b-1)^2$$

whose simultaneous solution is:

$$\hat{y}_{ij} = \frac{ay'_{i.}+by'_{.j}-y'_{..}}{(a-1)(b-1)} + \frac{(b-1)\left[ay'_{k.}+by'_{.j}-y'_{..}+(a-1)(b-1)^2\left(ay'_{i.}+by'_{.j}-y'_{..}\right)\right]}{\left[1-(a-1)^2(b-1)^4\right]}$$

$$\hat{y}_{kj} = \frac{ay'_{k.}+by'_{.j}-y'_{..}-(b-1)^2(a-1)\left[ay'_{i.}+by'_{.j}-y'_{..}\right]}{(a-1)(b-1)\left[1-(a-1)^2(b-1)^4\right]}$$

4.21. An industrial engineer is conducting an experiment on eye focus time. He is interested in the effect of the distance of the object from the eye on the focus time. Four different distances are of interest. He has five subjects available for the experiment. Because there may be differences among individuals, he decides to conduct the experiment in a randomized block design. The data obtained follow. Analyze the data from this experiment (use $\alpha = 0.05$) and draw appropriate conclusions.

	Subject				
Distance (ft)	1	2	3	4	5
4	10	6	6	6	6
6	7	6	6	1	6
8	5	3	3	2	5
10	6	4	4	2	3

Design Expert Output

Response: Focus Time
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	36.30	4	9.07			
Model	32.95	3	10.98	8.61	0.0025	significant
A	*32.95*	*3*	*10.98*	*8.61*	*0.0025*	
Residual	15.30	12	1.27			
Cor Total	84.55	19				

The Model F-value of 8.61 implies the model is significant. There is only a 0.25% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.13	R-Squared	0.6829
Mean	4.85	Adj R-Squared	0.6036
C.V.	23.28	Pred R-Squared	0.1192
PRESS	42.50	Adeq Precision	10.432

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	**Standard Error**
1-4	6.80	0.50
2-6	5.20	0.50
3-8	3.60	0.50
4-10	3.80	0.50

Treatment	**Mean Difference**	**DF**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	1.60	1	0.71	2.24	0.0448
1 vs 3	3.20	1	0.71	4.48	0.0008
1 vs 4	3.00	1	0.71	4.20	0.0012
2 vs 3	1.60	1	0.71	2.24	0.0448
2 vs 4	1.40	1	0.71	1.96	0.0736
3 vs 4	-0.20	1	0.71	-0.28	0.7842

Distance has a statistically significant effect on mean focus time.

4.22. The effect of five different ingredients (*A, B, C, D, E*) on reaction time of a chemical process is being studied. Each batch of new material is only large enough to permit five runs to be made. Furthermore, each run requires approximately 1 1/2 hours, so only five runs can be made in one day. The experimenter decides to run the experiment as a Latin square so that day and batch effects can be systematically controlled. She obtains the data that follow. Analyze the data from this experiment (use α = 0.05) and draw conclusions.

	Day				
Batch	1	2	3	4	5
1	*A*=8	*B*=7	*D*=1	*C*=7	*E*=3
2	*C*=11	*E*=2	*A*=7	*D*=3	*B*=8
3	*B*=4	*A*=9	*C*=10	*E*=1	*D*=5
4	*D*=6	*C*=8	*E*=6	*B*=6	*A*=10
5	*E*=4	*D*=2	*B*=3	*A*=8	*C*=8

The *Minitab* output below identifies the ingredients as having a significant effect on reaction time.

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Batch    random      5 1 2 3 4 5
Day      random      5 1 2 3 4 5
Catalyst  fixed      5 A B C D E

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Catalyst    4    141.440    141.440     35.360   11.31  0.000
Batch       4     15.440     15.440      3.860    1.23  0.348
Day         4     12.240     12.240      3.060    0.98  0.455
Error      12     37.520     37.520      3.127
Total      24    206.640
```

4.23. An industrial engineer is investigating the effect of four assembly methods (*A, B, C, D*) on the assembly time for a color television component. Four operators are selected for the study. Furthermore, the engineer knows that each assembly method produces such fatigue that the time required for the last assembly may be greater than the time required for the first, regardless of the method. That is, a trend develops in the required assembly time. To account for this source of variability, the engineer uses the Latin square design shown below. Analyze the data from this experiment (α = 0.05) draw appropriate conclusions.

Order of	Operator			
Assembly	1	2	3	4
1	*C*=10	*D*=14	*A*=7	*B*=8
2	*B*=7	*C*=18	*D*=11	*A*=8
3	*A*=5	*B*=10	*C*=11	*D*=9
4	*D*=10	*A*=10	*B*=12	*C*=14

The Minitab output below identifies assembly method as having a significant effect on assembly time.

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Order    random      4 1 2 3 4
Operator random      4 1 2 3 4
Method    fixed      4 A B C D


Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Method      3     72.500     72.500     24.167   13.81  0.004
Order       3     18.500     18.500      6.167    3.52  0.089
Operator    3     51.500     51.500     17.167    9.81  0.010
Error       6     10.500     10.500      1.750
Total      15    153.000
```

4.24. Consider the randomized complete block design in Problem 4.4. Assume that the days are random. Estimate the block variance component.

The block variance component is:

$$\hat{\sigma}_{\beta}^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{a} = \frac{[368.97 - 8.64]}{3} = 120.11$$

4.25. Consider the randomized complete block design in Problem 4.7. Assume that the coupons are random. Estimate the block variance component.

The block variance component is:

$$\hat{\sigma}_{\beta}^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{a} = \frac{[0.27 - 0.008889]}{4} = 0.06528$$

4.26. Consider the randomized complete block design in Problem 4.9. Assume that the trucks are random. Estimate the block variance component.

The block variance component is:

$$\hat{\sigma}_{\beta}^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{a} = \frac{[0.023 - 0.0005278]}{3} = 0.007491$$

4.27. Consider the randomized complete block design in Problem 4.11. Assume that the software projects that were used as blocks are random. Estimate the block variance component.

The block variance component is:

$$\hat{\sigma}_{\beta}^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{a} = \frac{[457500 - 111200]}{6} = 57716.67$$

4.28. Consider the gene expression experiment in Problem 4.12. Assume that the subjects used in this experiment are random. Estimate the block variance component

The block variance component is:

$$\hat{\sigma}_{\beta}^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{a} = \frac{[102300 - 73130.15]}{3} = 9723.28$$

4.29. Suppose that in Problem 4.22 the observation from batch 3 on day 4 is missing. Estimate the missing value and perform the analysis using this value.

y_{354} is missing. $\hat{y}_{354} = \dfrac{p\left[y'_{i..} + y'_{.j.} + y'_{..k}\right] - 2y'_{...}}{(p-2)(p-1)} = \dfrac{5[28+15+24] - 2(146)}{(3)(4)} = 3.58$

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Batch    random      5 1 2 3 4 5
Day      random      5 1 2 3 4 5
Catalyst  fixed      5 A B C D E

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Catalyst    4    128.676    128.676     32.169   11.25  0.000
Batch       4     16.092     16.092      4.023    1.41  0.290
Day         4      8.764      8.764      2.191    0.77  0.567
Error      12     34.317     34.317      2.860
Total      24    187.849
```

4.30. Consider a *p x p* Latin square with rows (α_i), columns (β_k), and treatments (τ_j) fixed. Obtain least squares estimates of the model parameters α_i, β_k, τ_j.

$$\mu : p^2\hat{\mu} + p\sum_{i=1}^{p}\hat{\alpha}_i + p\sum_{j=1}^{p}\hat{\tau}_j + p\sum_{k=1}^{p}\hat{\beta}_k = y_{...}$$

$$\alpha_i : p\hat{\mu} + p\hat{\alpha}_i + p\sum_{j=1}^{p}\hat{\tau}_j + p\sum_{k=1}^{p}\hat{\beta}_k = y_{i..} \quad , \; i = 1,2,...,p$$

$$\tau_j : p\hat{\mu} + p\sum_{i=1}^{p}\hat{\alpha}_i + p\hat{\tau}_j + p\sum_{k=1}^{p}\hat{\beta}_k = y_{.j.} , \; j = 1,2,...,p$$

$$\beta_k : p\hat{\mu} + p\sum_{i=1}^{p}\hat{\alpha}_i + p\sum_{j=1}^{p}\hat{\tau}_j + p\hat{\beta}_k = y_{..k} , \; k = 1,2,...,p$$

There are 3*p*+1 equations in 3*p*+1 unknowns. The rank of the system is 3*p*-2. Three side conditions are necessary. The usual conditions imposed are: $\sum_{i=1}^{p}\hat{\alpha}_i = \sum_{j=1}^{p}\hat{\tau}_j = \sum_{k=1}^{p}\hat{\beta}_k = 0$. The solution is then:

$$\hat{\mu} = \frac{y_{...}}{p^2} = \bar{y}_{...}$$
$$\hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...}, i = 1, 2, ..., p$$
$$\hat{\tau}_j = \bar{y}_{.j.} - \bar{y}_{...}, \; j = 1, 2, ..., p$$
$$\hat{\beta}_k = \bar{y}_{i..} - \bar{y}_{...}, k = 1, 2, ..., p$$

4.31. Derive the missing value formula (Equation 4.27) for the Latin square design.

$$SS_E = \sum\sum\sum y_{ijk}^2 - \sum\frac{y_{i..}^2}{p} - \sum\frac{y_{.j.}^2}{p} - \sum\frac{y_{..k}^2}{p} + 2\left(\frac{y_{...}^2}{p^2}\right)$$

Let y_{ijk} be missing. Then

$$SS_E = y_{ijk}^2 - \frac{\left(y'_{i..} + y_{ijk}\right)^2}{p} - \frac{\left(y'_{.j.} + y_{ijk}\right)^2}{p} - \frac{\left(y'_{..k} + y_{ijk}\right)^2}{p} + \frac{2\left(y'_{...} + y_{ijk}\right)^2}{p^2} + R$$

where R is all terms without y_{ijk}.. From $\frac{\partial SS_E}{\partial y_{ijk}} = 0$, we obtain:

$$y_{ijk}\,\frac{(p-1)(p-2)}{p^2} = \frac{p\left(y'_{i..} + y'_{.j.} + y'_{..k}\right) - 2y'_{...}}{p^2}, \text{ or } y_{ijk} = \frac{p\left(y'_{i..} + y'_{.j.} + y'_{..k}\right) - 2y'_{...}}{(p-1)(p-2)}$$

4.32. ***Designs involving several Latin squares***. [See Cochran and Cox (1957), John (1971).] The *p x p* Latin square contains only *p* observations for each treatment. To obtain more replications the experimenter may use several squares, say *n*. It is immaterial whether the squares used are the same are different. The appropriate model is

$$y_{ijkh} = \mu + \rho_h + \alpha_{i(h)} + \tau_j + \beta_{k(h)} + (\tau\rho)_{jh} + \varepsilon_{ijkh} \qquad \begin{cases} i = 1,2,\ldots,p \\ j = 1,2,\ldots,p \\ k = 1,2,\ldots,p \\ h = 1,2,\ldots,n \end{cases}$$

where y_{ijkh} is the observation on treatment *j* in row *i* and column *k* of the *h*th square. Note that $\alpha_{i(h)}$ and $\beta_{k(h)}$ are row and column effects in the *h*th square, and ρ_h is the effect of the *h*th square, and $(\tau\rho)_{jh}$ is the interaction between treatments and squares.

(a) Set up the normal equations for this model, and solve for estimates of the model parameters. Assume that appropriate side conditions on the parameters are $\sum_h \hat{\rho}_h = 0$, $\sum_i \hat{\alpha}_{i(h)} = 0$, and $\sum_k \hat{\beta}_{k(h)} = 0$ for each *h*, $\sum_j \hat{\tau}_j = 0$, $\sum_j (\hat{\tau\rho})_{jh} = 0$ for each *h*, and $\sum_h (\hat{\tau\rho})_{jh} = 0$ for each *j*.

$$\begin{aligned}
\hat{\mu} &= \bar{y}_{....} \\
\hat{\rho}_h &= \bar{y}_{...h} - \bar{y}_{....} \\
\hat{\tau}_j &= \bar{y}_{.j..} - \bar{y}_{....} \\
\hat{\alpha}_{i(h)} &= \bar{y}_{i..h} - \bar{y}_{...h} \\
\hat{\beta}_{k(h)} &= \bar{y}_{..kh} - \bar{y}_{...h} \\
\left(\widehat{\tau\rho}\right)_{jh} &= \bar{y}_{.j.h} - \bar{y}_{.j..} - \bar{y}_{...h} + \bar{y}_{....}
\end{aligned}$$

(b) Write down the analysis of variance table for this design.

Source	SS	DF
Treatments	$\sum \frac{y_{.j..}^2}{np} - \frac{y_{....}^2}{np^2}$	p-1
Squares	$\sum \frac{y_{...h}^2}{p^2} - \frac{y_{....}^2}{np^2}$	n-1
Treatment x Squares	$\sum \frac{y_{.j.h}^2}{p} - \frac{y_{....}^2}{np^2} - SS_{Treatments} - SS_{Squares}$	(p-1)(n-1)
Rows	$\sum \frac{y_{i..h}^2}{p} - \frac{y_{...h}^2}{np}$	n(p-1)
Columns	$\sum \frac{y_{..kh}^2}{p} - \frac{y_{...h}^2}{np}$	n(p-1)
Error	subtraction	n(p-1)(p-2)
Total	$\sum\sum\sum\sum y_{ijkh}^2 - \frac{y_{....}^2}{np^2}$	np^2-1

4.33. Discuss how the operating characteristics curves in the Appendix may be used with the Latin square design.

For the fixed effects model use:

$$\Phi^2 = \frac{\sum p\tau_j^2}{p\sigma^2} = \sum \frac{\tau_j^2}{\sigma^2}, \ \upsilon_1 = p-1 \quad \upsilon_2 = (p-2)(p-1)$$

For the random effects model use:

$$\lambda = \sqrt{1 + \frac{p\sigma_\tau^2}{\sigma^2}}, \ \upsilon_1 = p-1 \quad \upsilon_2 = (p-2)(p-1)$$

4.34. Suppose that in Problem 4.22 the data taken on day 5 were incorrectly analyzed and had to be discarded. Develop an appropriate analysis for the remaining data.

Two methods of analysis exist: (1) Use the general regression significance test, or (2) recognize that the design is a Youden square. The data can be analyzed as a balanced incomplete block design with $a = b = 5$, $r = k = 4$ and $\lambda = 3$. Using either approach will yield the same analysis of variance.

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Catalyst  fixed      5 A B C D E
Batch    random      5 1 2 3 4 5
Day      random      4 1 2 3 4
```

```
Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Catalyst    4    119.800    120.167     30.042    7.48  0.008
Batch       4     11.667     11.667      2.917    0.73  0.598
Day         3      6.950      6.950      2.317    0.58  0.646
Error       8     32.133     32.133      4.017
Total      19    170.550
```

4.35. The yield of a chemical process was measured using five batches of raw material, five acid concentrations, five standing times, (A, B, C, D, E) and five catalyst concentrations (α, β, γ, δ, ε). The Graeco-Latin square that follows was used. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

			Acid Concentration		
Batch	1	2	3	4	5
1	$A\alpha$=26	$B\beta$=16	$C\gamma$=19	$D\delta$=16	$E\varepsilon$=13
2	$B\gamma$=18	$C\delta$=21	$D\varepsilon$=18	$E\alpha$=11	$A\beta$=21
3	$C\varepsilon$=20	$D\alpha$=12	$E\beta$=16	$A\gamma$=25	$B\delta$=13
4	$D\beta$=15	$E\gamma$=15	$A\delta$=22	$B\varepsilon$=14	$C\alpha$=17
5	$E\delta$=10	$A\varepsilon$=24	$B\alpha$=17	$C\beta$=17	$D\gamma$=14

The Minitab output below identifies standing time as having a significant effect on yield.

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Time      fixed      5 A B C D E
Catalyst random      5 a b c d e
Batch    random      5 1 2 3 4 5
Acid     random      5 1 2 3 4 5

Analysis of Variance for Yield, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Time        4    342.800    342.800     85.700   14.65  0.001
Catalyst    4     12.000     12.000      3.000    0.51  0.729
Batch       4     10.000     10.000      2.500    0.43  0.785
Acid        4     24.400     24.400      6.100    1.04  0.443
Error       8     46.800     46.800      5.850
Total      24    436.000
```

4.36. Suppose that in Problem 4.23 the engineer suspects that the workplaces used by the four operators may represent an additional source of variation. A fourth factor, workplace (α, β, γ, δ) may be introduced and another experiment conducted, yielding the Graeco-Latin square that follows. Analyze the data from this experiment (use $\alpha = 0.05$) and draw conclusions.

Order of		Operator		
Assembly	1	2	3	4
1	$C\beta$=11	$B\gamma$=10	$D\delta$=14	$A\alpha$=8
2	$B\alpha$=8	$C\delta$=12	$A\gamma$=10	$D\beta$=12
3	$A\delta$=9	$D\alpha$=11	$B\beta$=7	$C\gamma$=15
4	$D\gamma$=9	$A\beta$=8	$C\alpha$=18	$B\delta$=6

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Method    fixed      4 A B C D
Order    random      4 1 2 3 4
Operator random      4 1 2 3 4
Workplac random      4 a b c d

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Method      3     95.500     95.500     31.833    3.47  0.167
Order       3      0.500      0.500      0.167    0.02  0.996
Operator    3     19.000     19.000      6.333    0.69  0.616
Workplac    3      7.500      7.500      2.500    0.27  0.843
Error       3     27.500     27.500      9.167
Total      15    150.000
```

Method and workplace do not have a significant effect on assembly time. However, there are only three degrees of freedom for error, so the test is not very sensitive.

4.37. Construct a 5 x 5 hypersquare for studying the effects of five factors. Exhibit the analysis of variance table for this design.

Three 5 x 5 orthogonal Latin Squares are:

ABCDE	$\alpha\beta\gamma\delta\varepsilon$	12345
BCDEA	$\gamma\delta\varepsilon\alpha\beta$	45123
CDEAB	$\varepsilon\alpha\beta\gamma\delta$	23451
DEABC	$\beta\gamma\delta\varepsilon\alpha$	51234
EABCD	$\delta\varepsilon\alpha\beta\gamma$	34512

Let rows = factor 1, columns = factor 2, Latin letters = factor 3, Greek letters = factor 4 and numbers = factor 5. The analysis of variance table is:

Source	SS	DF
Rows	$\frac{1}{5}\sum_{i=1}^{5} y_{i....}^2 - \frac{y_{.....}^2}{25}$	4
Columns	$\frac{1}{5}\sum_{m=1}^{5} y_{....m}^2 - \frac{y_{.....}^2}{25}$	4
Latin Letters	$\frac{1}{5}\sum_{j=1}^{5} y_{.j...}^2 - \frac{y_{.....}^2}{25}$	4
Greek Letters	$\frac{1}{5}\sum_{k=1}^{5} y_{..k..}^2 - \frac{y_{.....}^2}{25}$	4
Numbers	$\frac{1}{5}\sum_{l=1}^{5} y_{...l.}^2 - \frac{y_{.....}^2}{25}$	4
Error	SS_E by subtraction	4
Total	$\sum_{i=1}^{5}\sum_{j=1}^{5}\sum_{k=1}^{5}\sum_{l=1}^{5}\sum_{m=1}^{5} y_{ijklm}^2 - \frac{y_{.....}^2}{25}$	24

4.38. Consider the data in Problems 4.23 and 4.36. Suppressing the Greek letters in 4.36, analyze the data using the method developed in Problem 4.32.

Square 1 - Operator

Batch	1	2	3	4	Row Total
1	*C*=10	*D*=14	*A*=7	*B*=8	(39)
2	*B*=7	*C*=18	*D*=11	*A*=8	(44)
3	*A*=5	*B*=10	*C*=11	*D*=9	(35)
4	*D*=10	*A*=10	*B*=12	*C*=14	(46)
	(32)	(52)	(41)	(36)	164=$y_{\ldots 1}$

Square 2 - Operator

Batch	1	2	3	4	Row Total
1	*C*=11	*B*=10	*D*=14	*A*=8	(43)
2	*B*=8	*C*=12	*A*=10	*D*=12	(42)
3	*A*=9	*D*=11	*B*=7	*C*=15	(42)
4	*D*=9	*A*=8	*C*=18	*B*=6	(41)
	(37)	(41)	(49)	(41)	168=$y_{\ldots 2}$

Assembly Methods	Totals
A	$y_{.1..}$=65
B	$y_{.2..}$=68
C	$y_{.3..}$=109
D	$y_{.4..}$=90

Source	*SS*	*DF*	*MS*	F_0
Assembly Methods	159.25	3	53.08	14.00*
Squares	0.50	1	0.50	
A x *S*	8.75	3	2.92	0.77
Assembly Order (Rows)	19.00	6	3.17	
Operators (columns)	70.50	6	11.75	
Error	45.50	12	3.79	
Total	303.50	31		

Significant at 1%.

4.39. Consider the randomized block design with one missing value in Problem 4.19. Analyze this data by using the exact analysis of the missing value problem discussed in Section 4.1.4. Compare your results to the approximate analysis of these data given from Problem 4.19.

To simplify the calculations, the data in Problems 4.19 was transformed by multiplying by 10 and substracting 95.

$$
\begin{array}{llllllllllll}
\mu: & 15\hat{\mu} & +4\hat{\tau}_1 & +3\hat{\tau}_2 & +4\hat{\tau}_3 & +4\hat{\tau}_4 & +4\hat{\beta}_1 & +4\hat{\beta}_2 & +3\hat{\beta}_3 & +4\hat{\beta}_4 & =17 \\
\tau_1: & 4\hat{\mu} & +4\hat{\tau}_1 & & & & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =3 \\
\tau_2: & 3\hat{\mu} & & +3\hat{\tau}_2 & & & +\hat{\beta}_1 & +\hat{\beta}_2 & & +\hat{\beta}_4 & =1 \\
\tau_3: & 4\hat{\mu} & & & +4\hat{\tau}_3 & & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =-2 \\
\tau_4: & 4\hat{\mu} & & & & +4\hat{\tau}_4 & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =15 \\
\beta_1: & 4\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & +\hat{\tau}_4 & +4\hat{\beta}_1 & & & & =-4 \\
\beta_2: & 4\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & +\hat{\tau}_4 & & +4\hat{\beta}_2 & & & =-3 \\
\beta_3: & 3\hat{\mu} & +\hat{\tau}_1 & & +\hat{\tau}_3 & +\hat{\tau}_4 & & & +3\hat{\beta}_3 & & =6 \\
\beta_4: & 4\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & +\hat{\tau}_4 & & & & +4\hat{\beta}_4 & =18
\end{array}
$$

Applying the constraints $\sum \hat{\tau}_i = \sum \hat{\beta}_j = 0$, we obtain:

$$\hat{\mu} = \frac{41}{36},\ \hat{\tau}_1 = \frac{-14}{36},\ \hat{\tau}_2 = \frac{-24}{36},\ \hat{\tau}_3 = \frac{-59}{36}, \hat{\tau}_4 = \frac{94}{36},\ \hat{\beta}_1 = \frac{-77}{36},\ \hat{\beta}_2 = \frac{-68}{36},\ \hat{\beta}_3 = \frac{24}{36},\ \hat{\beta}_4 = \frac{121}{36}$$

$$R(\mu,\tau,\beta) = \hat{\mu}y_{..} + \sum_{i=1}^{4}\hat{\tau}_i y_{i.} + \sum_{j=1}^{4}\hat{\beta}_j y_{.j} = 138.78$$

With 7 degrees of freedom.

$$\sum\sum y_{ij}^2 = 145.00,\ SS_E = \sum\sum y_{ij}^2 - R(\mu,\tau,\beta) = 145.00 - 138.78 = 6.22$$

which is identical to SS_E obtained in the approximate analysis. In general, the SS_E in the exact and approximate analyses will be the same.

To test $H_o: \tau_i = 0$ the reduced model is $y_{ij} = \mu + \beta_j + \varepsilon_{ij}$. The normal equations used are:

$$
\begin{array}{lllllll}
\mu: & 15\hat{\mu} & +4\hat{\beta}_1 & +4\hat{\beta}_2 & +3\hat{\beta}_3 & +4\hat{\beta}_4 & =17 \\
\beta_1: & 4\hat{\mu} & +4\hat{\beta}_1 & & & & =-4 \\
\beta_2: & 4\hat{\mu} & & +4\hat{\beta}_2 & & & =-3 \\
\beta_3: & 3\hat{\mu} & & & +3\hat{\beta}_3 & & =6 \\
\beta_4: & 4\hat{\mu} & & & & +4\hat{\beta}_4 & =18
\end{array}
$$

Applying the constraint $\sum \hat{\beta}_j = 0$, we obtain:

$$\hat{\mu} = \frac{19}{16},\ \hat{\beta}_1 = \frac{-35}{16},\ \hat{\beta}_2 = \frac{-31}{16},\ \hat{\beta}_3 = \frac{13}{16},\ \hat{\beta}_4 = \frac{53}{16}.\ \text{Now}\ \ R(\mu,\beta) = \hat{\mu}y_{..} + \sum_{j=1}^{4}\hat{\beta}_j y_{.j} = 99.25$$

with 4 degrees of freedom.

$$R(\tau|\mu,\beta) = R(\mu,\tau,\beta) - R(\mu,\beta) = 138.78 - 99.25 = 39.53 = SS_{Treatments}$$

with 7-4=3 degrees of freedom. $R(\tau|\mu,\beta)$ is used to test H$_o$: $\tau_i = 0$.

The sum of squares for blocks is found from the reduced model $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$. The normal equations used are:

Model Restricted to $\beta_j = 0$:

$$
\begin{array}{llllllll}
\mu: & 15\hat{\mu} & +4\hat{\tau}_1 & +3\hat{\tau}_2 & +4\hat{\tau}_3 & +4\hat{\tau}_4 & =17 \\
\tau_1: & 4\hat{\mu} & +4\hat{\tau}_1 & & & & =3 \\
\tau_2: & 3\hat{\mu} & & +3\hat{\tau}_2 & & & =1 \\
\tau_3: & 4\hat{\mu} & & & +4\hat{\tau}_3 & & =-2 \\
\tau_4: & 4\hat{\mu} & & & & +4\hat{\tau}_4 & =15
\end{array}
$$

Applying the constraint $\sum \hat{\tau}_i = 0$, we obtain:

$$\hat{\mu} = \frac{13}{12},\ \hat{\tau}_1 = \frac{-4}{12},\ \hat{\tau}_2 = \frac{-9}{12},\ \hat{\tau}_3 = \frac{-19}{12},\ \hat{\tau}_4 = \frac{32}{12}$$

$$R(\mu,\tau) = \hat{\mu}y_{..} + \sum_{i=1}^{4} \hat{\tau}_i y_{i.} = 59.83$$

with 4 degrees of freedom.

$$R(\beta|\mu,\tau) = R(\mu,\tau,\beta) - R(\mu,\tau) = 138.78 - 59.83 = 78.95 = SS_{Blocks}$$

with 7-4=3 degrees of freedom.

Source	*DF*	*SS*(exact)	*SS*(approximate)
Tips	3	39.53	39.98
Blocks	3	78.95	79.53
Error	8	6.22	6.22
Total	14	125.74	125.73

Note that for the exact analysis, $SS_T \neq SS_{Tips} + SS_{Blocks} + SS_E$.

4.40. An engineer is studying the mileage performance characteristics of five types of gasoline additives. In the road test he wishes to use cars as blocks; however, because of a time constraint, he must use an incomplete block design. He runs the balanced design with the five blocks that follow. Analyze the data from this experiment (use α = 0.05) and draw conclusions.

	Car				
Additive	1	2	3	4	5
1		17	14	13	12
2	14	14		13	10
3	12		13	12	9
4	13	11	11	12	
5	11	12	10		8

There are several computer software packages that can analyze the incomplete block designs discussed in this chapter. The *Minitab* General Linear Model procedure is a widely available package with this capability. The output from this routine follows. The adjusted sums of squares are the appropriate sums of squares to use for testing the difference between the means of the gasoline additives. The gasoline additives have a significant effect on the mileage.

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Additive  fixed      5 1 2 3 4 5
Car      random      5 1 2 3 4 5

Analysis of Variance for Mileage, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Additive    4    31.7000    35.7333     8.9333    9.81  0.001
Car         4    35.2333    35.2333     8.8083    9.67  0.001
Error      11    10.0167    10.0167     0.9106
Total      19    76.9500
```

4.41. Construct a set of orthogonal contrasts for the data in Problem 4.40. Compute the sum of squares for each contrast.

One possible set of orthogonal contrasts is:

$$H_0 : \mu_4 + \mu_5 = \mu_1 + \mu_2 \qquad (1)$$
$$H_0 : \mu_1 = \mu_2 \qquad (2)$$
$$H_0 : \mu_4 = \mu_5 \qquad (3)$$
$$H_0 : 4\mu_3 = \mu_4 + \mu_5 + \mu_1 + \mu_2 \qquad (4)$$

The sums of squares and *F*-tests are:

Brand ->	1	2	3	4	5			
Q_i	33/4	11/4	-3/4	-14/4	-27/4	$\sum c_i Q_i$	*SS*	F_0
(1)	-1	-1	0	1	1	-85/4	30.10	33.06
(2)	1	-1	0	0	0	22/4	4.03	4.426
(3)	0	0	0	-1	1	-13/4	1.41	1.55
(4)	-1	-1	4	-1	-1	-15/4	0.19	0.21

Contrasts (1) and (2) are significant at the 1% and 5% levels, respectively.

4.42. Seven different hardwood concentrations are being studied to determine their effect on the strength of the paper produced. However the pilot plant can only produce three runs each day. As days may differ, the analyst uses the balanced incomplete block design that follows. Analyze this experiment (use $\alpha = 0.05$) and draw conclusions.

Hardwood	Days						
Concentration (%)	1	2	3	4	5	6	7
2	114				120		117
4	126	120				119	
6		137	117				134
8	141		129	149			
10		145		150	143		
12			120		118	123	
14				136		130	127

There are several computer software packages that can analyze the incomplete block designs discussed in this chapter. The Minitab General Linear Model procedure is a widely available package with this capability. The output from this routine for Problem 4.35 follows. The adjusted sums of squares are the appropriate sums of squares to use for testing the difference between the means of the hardwood concentrations.

Minitab Output

General Linear Model

```
Factor     Type Levels Values
Concentr  fixed      7  2  4  6  8 10 12 14
Days     random      7 1 2 3 4 5 6 7

Analysis of Variance for Strength, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Concentr    6    2037.62    1317.43     219.57   10.42  0.002
Days        6     394.10     394.10      65.68    3.12  0.070
Error       8     168.57     168.57      21.07
Total      20    2600.29
```

4.43. Analyze the data in Example 4.5 using the general regression significance test.

$$
\begin{array}{lllllllllll}
\mu: & 12\hat{\mu} & +3\hat{\tau}_1 & +3\hat{\tau}_2 & +3\hat{\tau}_3 & +3\hat{\tau}_4 & +3\hat{\beta}_1 & +3\hat{\beta}_2 & +3\hat{\beta}_3 & +3\hat{\beta}_4 & =870 \\
\tau_1: & 3\hat{\mu} & +3\hat{\tau}_1 & & & & +\hat{\beta}_1 & +\hat{\beta}_2 & & +\hat{\beta}_4 & =218 \\
\tau_2: & 3\hat{\mu} & & +3\hat{\tau}_2 & & & & +\hat{\beta}_2 & +\hat{\beta}_3 & +\hat{\beta}_4 & =214 \\
\tau_3: & 3\hat{\mu} & & & +3\hat{\tau}_3 & & +\hat{\beta}_1 & +\hat{\beta}_2 & +\hat{\beta}_3 & & =216 \\
\tau_4: & 3\hat{\mu} & & & & +3\hat{\tau}_4 & +\hat{\beta}_1 & & +\hat{\beta}_3 & +\hat{\beta}_4 & =222 \\
\beta_1: & 3\hat{\mu} & +\hat{\tau}_1 & & +\hat{\tau}_3 & +\hat{\tau}_4 & +3\hat{\beta}_1 & & & & =221 \\
\beta_2: & 3\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & +\hat{\tau}_3 & & & +3\hat{\beta}_2 & & & =224 \\
\beta_3: & 3\hat{\mu} & & +\hat{\tau}_2 & +\hat{\tau}_3 & +\hat{\tau}_4 & & & +3\hat{\beta}_3 & & =207 \\
\beta_4: & 3\hat{\mu} & +\hat{\tau}_1 & +\hat{\tau}_2 & & +\hat{\tau}_4 & & & & +3\hat{\beta}_4 & =218
\end{array}
$$

Applying the constraints $\sum \hat{\tau}_i = \sum \hat{\beta}_j = 0$, we obtain:

$$\hat{\mu} = 870/12,\ \hat{\tau}_1 = -9/8,\ \hat{\tau}_2 = -7/8,\ \hat{\tau}_3 = -4/8,\ \hat{\tau}_4 = 20/8,$$

$$\hat{\beta}_1 = 7/8,\ \hat{\beta}_2 = 24/8,\ \hat{\beta}_3 = -31/8,\ \hat{\beta}_4 = 0/8$$

$$R(\mu,\tau,\beta) = \hat{\mu} y_{..} + \sum_{i=1}^{4} \hat{\tau}_i y_{i.} + \sum_{j=1}^{4} \hat{\beta}_j y_{.j} = 63{,}152.75$$

with 7 degrees of freedom.

$$\sum \sum y_{ij}^2 = 63{,}156.00$$

$$SS_E = \sum \sum y_{ij}^2 - R(\mu,\tau,\beta) = 63156.00 - 63152.75 = 3.25 .$$

To test H$_o$: $\tau_i = 0$ the reduced model is $y_{ij} = \mu + \beta_j + \varepsilon_{ij}$. The normal equations used are:

$$\begin{array}{llllllll}
\mu: & 12\hat{\mu} & +3\hat{\beta}_1 & +3\hat{\beta}_2 & +3\hat{\beta}_3 & +3\hat{\beta}_4 & = 870 \\
\beta_1: & 3\hat{\mu} & +3\hat{\beta}_1 & & & & = 221 \\
\beta_2: & 3\hat{\mu} & & +3\hat{\beta}_2 & & & = 224 \\
\beta_3: & 3\hat{\mu} & & & +3\hat{\beta}_3 & & = 207 \\
\beta_4: & 3\hat{\mu} & & & & +3\hat{\beta}_4 & = 218
\end{array}$$

Applying the constraint $\sum \hat{\beta}_j = 0$, we obtain:

$$\hat{\mu} = \frac{870}{12},\ \hat{\beta}_1 = \frac{7}{6},\ \hat{\beta}_2 = \frac{13}{6},\ \hat{\beta}_3 = \frac{-21}{6},\ \hat{\beta}_4 = \frac{1}{6}$$

$$R(\mu,\beta) = \hat{\mu} y_{..} + \sum_{j=1}^{4} \hat{\beta}_j y_{.j} = 63{,}130.00$$

with 4 degrees of freedom.

$$R(\tau|\mu,\beta) = R(\mu,\tau,\beta) - R(\mu,\beta) = 63152.75 - 63130.00 = 22.75 = SS_{Treatments}$$

with 7 – 4 = 3 degrees of freedom. $R(\tau|\mu,\beta)$ is used to test H$_o$: $\tau_i = 0$.

The sum of squares for blocks is found from the reduced model $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$. The normal equations used are:

Model Restricted to $\beta_j = 0$:

$$\begin{array}{llllllll}
\mu: & 12\hat{\mu} & +3\hat{\tau}_1 & +3\hat{\tau}_2 & +3\hat{\tau}_3 & +3\hat{\tau}_4 & = 870 \\
\tau_1: & 3\hat{\mu} & +3\hat{\tau}_1 & & & & = 218 \\
\tau_2: & 3\hat{\mu} & & +3\hat{\tau}_2 & & & = 214 \\
\tau_3: & 3\hat{\mu} & & & +3\hat{\tau}_3 & & = 216 \\
\tau_4: & 3\hat{\mu} & & & & +3\hat{\tau}_4 & = 222
\end{array}$$

The sum of squares for blocks is found as in Example 4.5. We may use the method shown above to find an adjusted sum of squares for blocks from the reduced model, $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$.

4.44. Prove that $\dfrac{k\sum_{i=1}^{a} Q_i^2}{(\lambda a)}$ is the adjusted sum of squares for treatments in a BIBD.

We may use the general regression significance test to derive the computational formula for the adjusted treatment sum of squares. We will need the following:

$$\hat{\tau}_i = \frac{kQ_i}{(\lambda a)},\ kQ_i = ky_{i.} - \sum_{i=1}^{b} n_{ij} y_{.j}$$

$$R(\mu,\tau,\beta) = \hat{\mu} y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} + \sum_{j=1}^{b} \hat{\beta}_j y_{.j}$$

and the sum of squares we need is:

$$R(\tau|\mu,\beta) = \hat{\mu} y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} + \sum_{j=1}^{b} \hat{\beta}_j y_{.j} - \sum_{j=1}^{b} \frac{y_{.j}^2}{k}$$

The normal equation for β is, from equation (4.35),

$$\beta : k\hat{\mu} + \sum_{i=1}^{a} n_{ij}\hat{\tau}_i + k\hat{\beta}_j = y_{.j}$$

and from this we have:

$$ky_{.j}\hat{\beta}_j = y_{.j}^2 - ky_{.j}\hat{\mu} - y_{.j}\sum_{i=1}^{a} n_{ij}\hat{\tau}_i$$

therefore,

$$R(\tau|\mu,\beta) = \hat{\mu} y_{..} + \sum_{i=1}^{a} \hat{\tau}_i y_{i.} + \sum_{j=1}^{b}\left[\frac{y_{.j}^2}{k} - \frac{k\hat{\mu} y_{.j}}{k} - \frac{y_{.j}\sum_{i=1}^{a} n_{ij}\hat{\tau}_i}{k} - \frac{y_{.j}^2}{k}\right]$$

$$R(\tau|\mu,\beta) = \sum_{i=1}^{a} \hat{\tau}_i\left(y_{i.} - \frac{1}{k}\sum_{i=1}^{a} n_{ij} y_{.j}\right) = \sum_{i=1}^{a} Q_i\left(\frac{kQ_i}{\lambda a}\right) = k\sum_{i=1}^{a}\left(\frac{Q_i^2}{\lambda a}\right) \equiv SS_{Treatments(adjusted)}$$

4.45. An experimenter wishes to compare four treatments in blocks of two runs. Find a BIBD for this experiment with six blocks.

Treatment	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
1	X	X	X			
2	X			X	X	
3		X		X		X
4			X		X	X

Note that the design is formed by taking all combinations of the 4 treatments 2 at a time. The parameters of the design are $\lambda = 1$, $a = 4$, $b = 6$, $k = 3$, and $r = 2$

4.46. An experimenter wishes to compare eight treatments in blocks of four runs. Find a BIBD with 14 blocks and $\lambda = 3$.

The design has parameters $a = 8$, $b = 14$, $\lambda = 3$, $r = 2$ and $k = 4$. It may be generated from a 2^3 factorial design confounded in two blocks of four observations each, with each main effect and interaction successively confounded (7 replications) forming the 14 blocks. The design is discussed by John (1971, pg. 222) and Cochran and Cox (1957, pg. 473). The design follows:

Blocks	1=(*I*)	2=*a*	3=*b*	4=*ab*	5=*c*	6=*ac*	7=*bc*	8=*abc*
1	X		X		X		X	
2		X		X		X		X
3	X		X			X		X
4		X		X	X		X	
5	X	X			X	X		
6			X	X			X	X
7	X	X					X	X
8			X	X	X	X		
9	X	X	X	X				
10					X	X	X	X
11	X			X		X	X	
12		X	X		X			X
13	X			X	X			X
14		X	X			X	X	

4.47. Perform the interblock analysis for the design in Problem 4.40.

The interblock analysis for Problem 4.33 uses $\hat{\sigma}^2 = 0.91$ and $\hat{\sigma}_\beta^2 = 2.63$. A summary of the interblock, intrablock and combined estimates is:

Parameter	Intrablock	Interblock
τ_1	2.20	-1.80
τ_2	0.73	0.20
τ_3	-0.20	-5.80
τ_4	-0.93	9.20
τ_5	-1.80	-1.80

4.48. Perform the interblock analysis for the design in Problem 4.42.

The interblock analysis for Pproblem 4.42 uses $\hat{\sigma}^2 = 21.07$ and

$$\sigma_\beta^2 = \frac{\left[MS_{Blocks(adj)} - MS_E\right](b-1)}{a(r-1)} = \frac{[65.68 - 21.07](6)}{7(2)} = 19.12 .$$

A summary of the interblock, intrablock, and combined estimates is give below

Parameter	Intrablock	Interblock	Combined
τ_1	-12.43	-11.79	-12.38
τ_2	-8.57	-4.29	-7.92
τ_3	2.57	-8.79	1.76
τ_4	10.71	9.21	10.61
τ_5	13.71	21.21	14.67
τ_6	-5.14	-22.29	-6.36
τ_7	-0.86	10.71	-0.03

4.49. Verify that a BIBD with the parameters a = 8, r = 8, k = 4, and b = 16 does not exist.

These conditions imply that $\lambda = \frac{r(k-1)}{a-1} = \frac{8(3)}{7} = \frac{24}{7}$, which is not an integer, so a balanced design with these parameters cannot exist.

4.50. Show that the variance of the intra block estimators $\{\hat{\tau}_i\}$ is $\frac{k((a-1))\sigma^2}{(\lambda a^2)}$.

Note that $\hat{\tau}_i = \frac{kQ_i}{(\lambda a)}$, and $Q_i = y_{i.} - \frac{1}{k}\sum_{j=1}^{b} n_{ij} y_{.j}$, and $kQ_i = ky_{i.} - \sum_{j=1}^{b} n_{ij} y_{.j} = (k-1)y_{i.} - \left(\sum_{j=1}^{b} n_{ij} y_{.j} - y_{i.}\right)$

$y_{i.}$ contains r observations, and the quantity in the parenthesis is the sum of $r(k\text{-}1)$ observations, not including treatment i. Therefore,

$$V(kQ_i) = k^2 V(Q_i) = r(k-1)^2\sigma^2 + r(k-1)\sigma^2$$

or

$$V(Q_i) = \frac{1}{k^2}\left[r(k-1)\sigma^2\{(k-1)+1\}\right] = \frac{r(k-1)\sigma^2}{k}$$

To find $V(\hat{\tau}_i)$, note that:

$$V(\hat{\tau}_i) = \left(\frac{k}{\lambda a}\right)^2 V(Q)_i = \left(\frac{k}{\lambda a}\right)^2 \frac{r(k-1)}{k}\sigma^2 = \frac{kr(k-1)}{(\lambda a)^2}\sigma^2$$

However, since $\lambda(a-1) = r(k-1)$, we have:

$$V(\hat{\tau}_i)=\frac{k(a-1)}{\lambda a^2}\sigma^2$$

Furthermore, the $\{\hat{\tau}_i\}$ are not independent, this is required to show that $V(\hat{\tau}_i-\hat{\tau}_j)=\frac{2k}{\lambda a}\sigma^2$

4.51. ***Extended incomplete block designs.*** Occasionally the block size obeys the relationship $a < k < 2a$. An extended incomplete block design consists of a single replicate or each treatment in each block along with an incomplete block design with $k^* = k\text{-}a$. In the balanced case, the incomplete block design will have parameters $k^* = k\text{-}a$, $r^* = r\text{-}b$, and λ^*. Write out the statistical analysis. (Hint: In the extended incomplete block design, we have $\lambda = 2r\text{-}b+\lambda^*$.)

As an example of an extended incomplete block design, suppose we have a=5 treatments, b=5 blocks and k=9. A design could be found by running all five treatments in each block, plus a block from the balanced incomplete block design with $k^* = k\text{-}a$=9-5=4 and λ^*=3. The design is:

Block	Complete Treatment	Incomplete Treatment
1	1,2,3,4,5	2,3,4,5
2	1,2,3,4,5	1,2,4,5
3	1,2,3,4,5	1,3,4,5
4	1,2,3,4,5	1,2,3,4
5	1,2,3,4,5	1,2,3,5

Note that r=9, since the augmenting incomplete block design has r*=4, and $r= r^* + b$ = 4+5=9, and $\lambda = 2r\text{-}b+\lambda^*$=18-5+3=16. Since some treatments are repeated in each block it is possible to compute an error sum of squares between repeat observations. The difference between this and the residual sum of squares is due to interaction. The analysis of variance table is shown below:

Source	*SS*	*DF*
Treatments (adjusted)	$k\sum\frac{Q_i^2}{a\lambda}$	a-1
Blocks	$\sum\frac{y_{.j}^2}{k}-\frac{y_{..}^2}{N}$	b-1
Interaction	Subtraction	(a-1)(b-1)
Error	[SS between repeat observations]	b(k-a)
Total	$\sum\sum y_{ij}^2-\frac{y_{..}^2}{N}$	N-1

4.52. Suppose that a single-factor experiment with five levels of the factor has been conducted. There are three replicates and the experiment has been conducted as a complete randomized design. If the experiment had been conducted in blocks, the pure error degrees of freedom would be reduced by (choose the correct answer):

(c) 2

4.53. Physics graduate student Laura Van Ertia has conducted a complete randomized design with a single factor, hoping to solve the mystery of the unified theory and complete her dissertation. The results of this experiment are summarized in the following ANOVA display:

Source	*DF*	*SS*	*MS*	*F*
Factor	-	-	14.18	-
Error	-	37.75	-	
Total	23	108.63		

The completed ANOVA is as follows:

Source	*DF*	*SS*	*MS*	*F*	*P*
Factor	5	70.88	14.18	6.76	0.00104
Error	18	37.75	2.10		
Total	23	108.63			

Answer the following questions about this experiment.

(a) The sum of squares for the factor is <u>70.88</u>.

(b) The number of degrees of freedom for the single factor in the experiment is <u>5</u>.

(c) The number of degrees of freedom for the error is <u>18</u>.

(d) The mean square for error is <u>2.10</u>.

(e) The value of the test statistic is <u>6.67</u>.

(f) If the significance level is 0.05, your conclusions are not to reject the null hypothesis. <u>No</u>.

(g) An upper bound on the *P*-value for the test statistic is <u>0.001</u>.

(h) A lower bound on the *P*-value for the test statistic is <u>0.0001</u>.

(i) Laura used <u>6</u> levels of the factor in this experiment.

(j) Laura replicated this experiment <u>4</u> times.

(k) Suppose that Laura had actually conducted this experiment as a random complete block design and the sum of squares for the blocks was 12. Reconstruct the ANOVA display above to reflect this new situation. How much has the blocking reduced the estimate of the experimental error?

Source	*DF*	*SS*	*MS*	*F*	*P*
Block	3	12.00	4.00		
Factor	5	70.88	14.18	9.91	0.00011
Error	18	25.75	1.43		
Total	23	108.63			

The blocking reduced the SS_{error} by 12 and the MS_{error} by 0.67 (32%).

4.54. Consider the direct mail marketing experiment in Problem 4.8. suppose that this experiment has been run as a complete randomized design, ignoring potential regional differences, but that exactly the same data was obtained. Reanalyze the experiment under this new assumption. What difference would ignoring the blocking have on the results and conclusions?

The solution for Problem 4.8 used a square root transformation, so the solution below also includes this same transformation. The results below are similar to Problem 4.8 in that the the difference in designs is statistically significant; however, the *F* value changed from 60.46 to only 7.02. The corresponding *P* value increased from 0.0001 to 0.0145.

Response: Number of responses 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	60.73	2	30.37	7.02	0.0145	significant
A-Design	*60.73*	*2*	*30.37*	*7.02*	*0.0145*	
Pure Error	38.90	9	4.32			
Cor Total	99.64	11				

The Model F-value of 7.02 implies the model is significant. There is only
a 1.45% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.08	R-Squared	0.6095
Mean	18.52	Adj R-Squared	0.5228
C.V. %	11.23	Pred R-Squared	0.3058
PRESS	69.16	Adeq Precision	4.803

Treatment Means (Adjusted, If Necessary)

	Estimated Mean	Standard Error
1-1	17.17	1.04
2-2	21.69	1.04
3-3	16.69	1.04

Treatment	Mean Difference	DF	Standard Error	t for H_0 Coeff=0	Prob > |t|
1 vs 2	-4.52	1	1.47	-3.07	0.0133
1 vs 3	0.48	1	1.47	0.33	0.7525
2 vs 3	4.99	1	1.47	3.40	0.0079

Chapter 5
Introduction to Factorial Designs
Solutions

5.1. The following output was obtained from a computer program that performed a two-factor ANOVA on a factorial experiment.

Two-way ANOVA: y versus A, B

Source	DF	SS	MS	F	P
A	1	0.322	?	?	?
B	?	80.554	40.2771	4.59	?
Interaction	?	?	?	?	?
Error	12	105.327	8.7773		
Total	17	231.551			

(a) Fill in the blanks in the ANOVA table. You can use bounds on the *P*-values.

Two-way ANOVA: y versus A, B

Source	DF	SS	MS	F	P
A	1	0.322	0.3220	0.04	0.8513
B	2	80.554	40.2771	4.59	0.0331
Interaction	2	45.348	22.6740	2.58	0.1167
Error	12	105.327	8.7773		
Total	17	231.551			

(b) How many levels were used for factor *B*?

3 levels.

(c) How many replicates of the experiment were performed?

3 replicates.

(d) What conclusions would you draw about this experiment?

Only factor *B* is significant; factor *A* and the two-factor interaction are not significant.

5.2. The following output was obtained from a computer program that performed a two-factor ANOVA on a factorial experiment.

Two-way ANOVA: y versus A, B

Source	DF	SS	MS	F	P
A	1	?	0.0002	?	?
B	?	180.378	?	?	?
Interaction	3	8.479	?	?	0.932
Error	8	158.797	?		
Total	15	347.653			

(e) Fill in the blanks in the ANOVA table. You can use bounds on the *P*-values.

Two-way ANOVA: y versus A, B

Source	DF	SS	MS	F	P
A	1	0.0002	0.0002	0.00001	0.998
B	3	180.378	60.1260	3.02907	0.093
Interaction	3	8.479	2.8263	0.14239	0.932
Error	8	158.797	19.8496		
Total	15	347.653			

(a) How many levels were used for factor *B*?

4 levels.

(b) How many replicates of the experiment were performed?

2 replicates.

(c) What conclusions would you draw about this experiment?

Factor *B* is moderately significant with a *P*-value of 0.93. Factor *A* and the two-factor interaction are not significant.

5.3. The yield of a chemical process is being studied. The two most important variables are thought to be the pressure and the temperature. Three levels of each factor are selected, and a factorial experiment with two replicates is performed. The yield data follow:

	Pressure		
Temperature	200	215	230
150	90.4	90.7	90.2
	90.2	90.6	90.4
160	90.1	90.5	89.9
	90.3	90.6	90.1
170	90.5	90.8	90.4
	90.7	90.9	90.1

(d) Analyze the data and draw conclusions. Use $\alpha = 0.05$.

Both pressure (*A*) and temperature (*B*) are significant, the interaction is not.

Design Expert Output

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

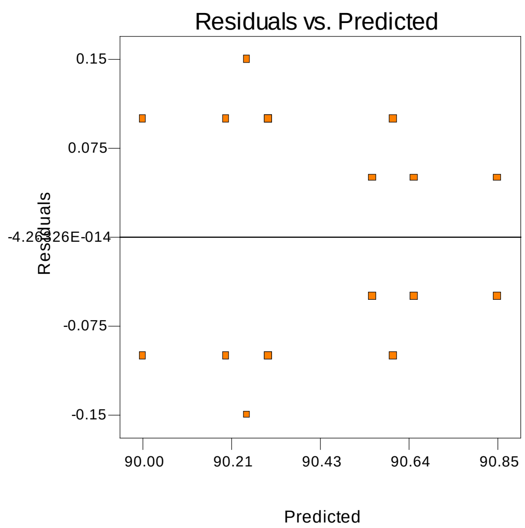
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1.14	8	0.14	8.00	0.0026	significant
A	*0.77*	*2*	*0.38*	*21.59*	*0.0004*	
B	*0.30*	*2*	*0.15*	*8.47*	*0.0085*	
AB	*0.069*	*4*	*0.017*	*0.97*	*0.4700*	
Residual	0.16	9	0.018			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.16*	*9*	*0.018*			
Cor Total	1.30	17				

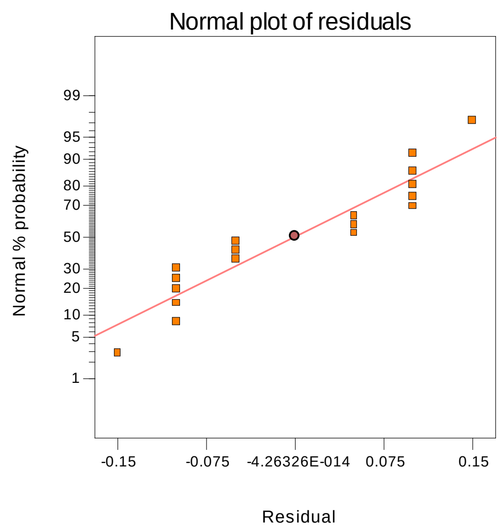
The Model F-value of 8.00 implies the model is significant. There is only a 0.26% chance that a "Model F-Value" this large could occur due to noise.

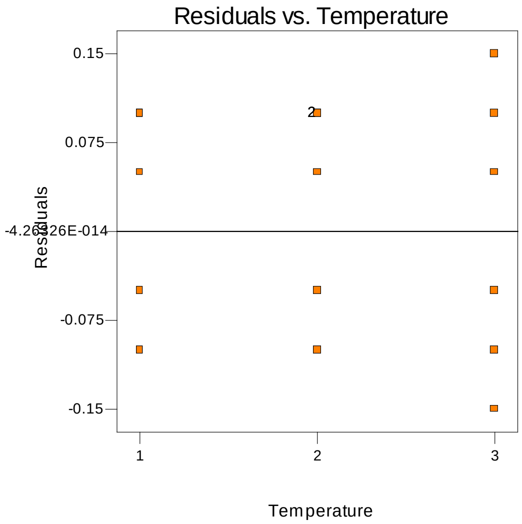
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

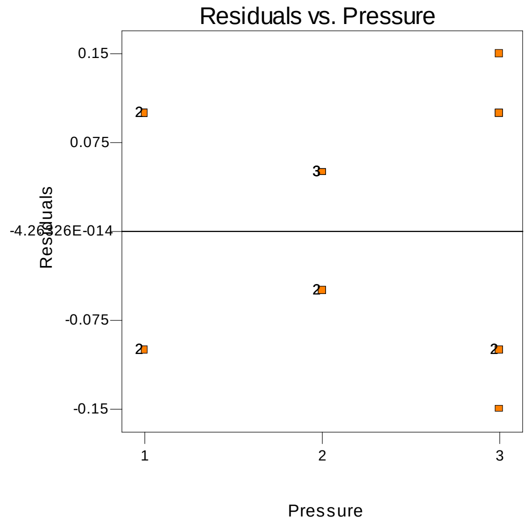
(b) Prepare appropriate residual plots and comment on the model's adequacy.

The residual plots show no serious deviations from the assumptions.

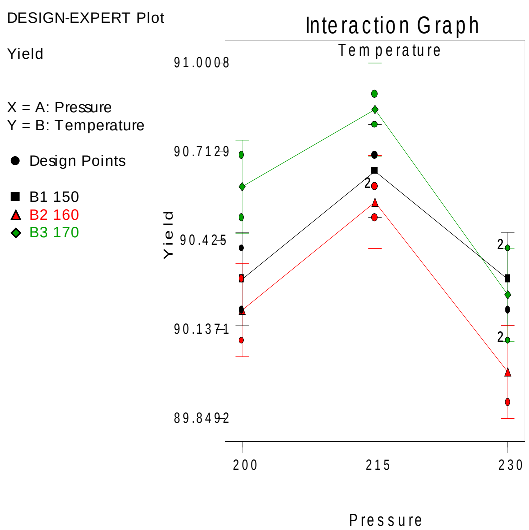








(c) Under what conditions would you operate this process?



Set pressure at 215 and Temperature at the high level, 170 degrees C, as this gives the highest yield.

The standard analysis of variance treats all design factors as if they were qualitative. In this case, both factors are quantitative, so some further analysis can be performed. In Section 5.5, we show how response curves and surfaces can be fit to the data from a factorial experiment with at least one quantative factor. Since both factors in this problem are quantitative and have three levels, we can fit linear and quadratic effects of both temperature and pressure, exactly as in Example 5.5 in the text. The Design-Expert output, including the response surface plots, now follows.

Design Expert Output

Response:Surface Finish
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	1.13	5	0.23	16.18	< 0.0001	significant
A	*0.10*	*1*	*0.10*	*7.22*	*0.0198*	
B	*0.067*	*1*	*0.067*	*4.83*	*0.0483*	
A2	*0.67*	*1*	*0.67*	*47.74*	*< 0.0001*	
B2	*0.23*	*1*	*0.23*	*16.72*	*0.0015*	
AB	*0.061*	*1*	*0.061*	*4.38*	*0.0582*	
Residual	0.17	12	0.014			
Lack of Fit	*7.639E-003*	*3*	*2.546E-003*	*0.14*	*0.9314*	*not significant*
Pure Error	*0.16*	*9*	*0.018*			
Cor Total	1.30	17				

The Model F-value of 16.18 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 A, B, A^2, B^2 are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy),
model reduction may improve your model.

Std. Dev.	0.12	R-Squared	0.8708
Mean	90.41	Adj R-Squared	0.8170
C.V.	0.13	Pred R-Squared	0.6794
PRESS	0.42	Adeq Precision	11.968

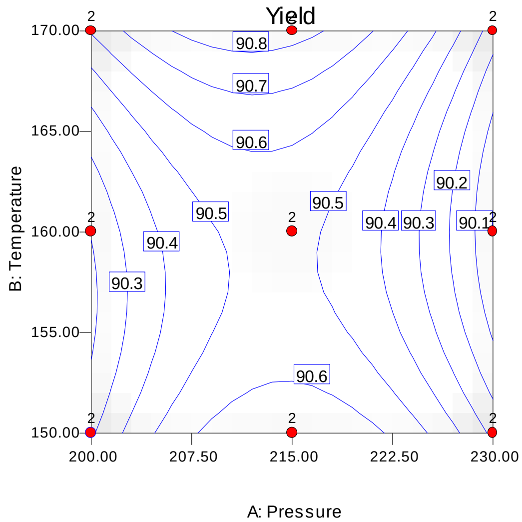
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	90.52	1	0.062	90.39	90.66	
A-Pressure	-0.092	1	0.034	-0.17	-0.017	1.00
B-Temperature	0.075	1	0.034	6.594E-004	0.15	1.00
A^2	-0.41	1	0.059	-0.54	-0.28	1.00
B^2	0.24	1	0.059	0.11	0.37	1.00
AB	-0.087	1	0.042	-0.18	3.548E-003	1.00

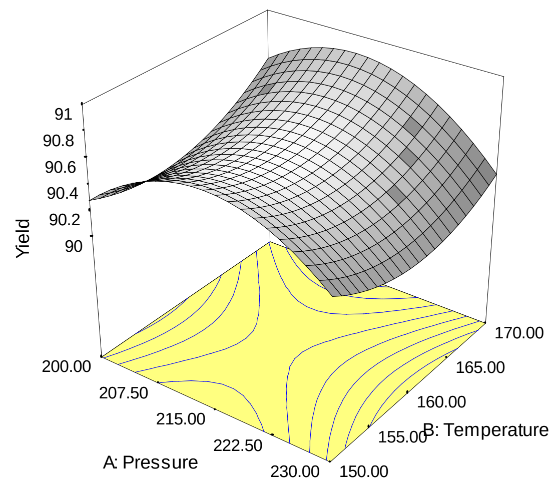
Final Equation in Terms of Coded Factors:

Yield	=
+90.52	
-0.092	* A
+0.075	* B
-0.41	* A^2
+0.24	* B^2
-0.087	* A * B

Final Equation in Terms of Actual Factors:

Yield	=
+48.54630	
+0.86759	* Pressure
-0.64042	* Temperature
-1.81481E-003	* Pressure^2
+2.41667E-003	* Temperature^2
-5.83333E-004	* Pressure * Temperature





5.4. An engineer suspects that the surface finish of a metal part is influenced by the feed rate and the depth of cut. She selects three feed rates and four depths of cut. She then conducts a factorial experiment and obtains the following data:

	Depth of Cut (in)			
Feed Rate (in/min)	0.15	0.18	0.20	0.25
	74	79	82	99
0.20	64	68	88	104
	60	73	92	96
	92	98	99	104
0.25	86	104	108	110
	88	88	95	99
	99	104	108	114
0.30	98	99	110	111
	102	95	99	107

(a) Analyze the data and draw conclusions. Use $\alpha = 0.05$.

The depth (A) and feed rate (B) are significant, as is the interaction (AB).

Design Expert Output

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

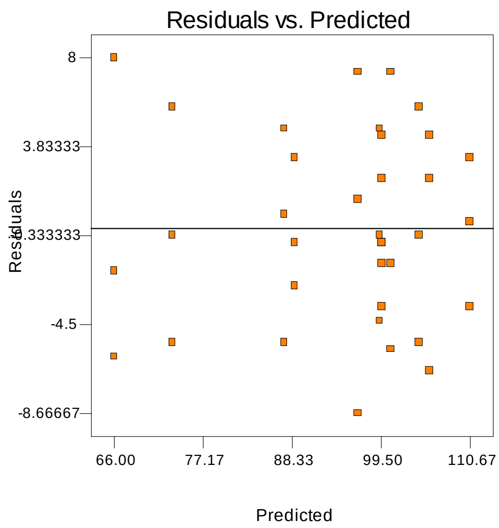
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	5842.67	11	531.15	18.49	< 0.0001	significant
A-Depth	*2125.11*	*3*	*708.37*	*24.66*	*< 0.0001*	
B-Feed	*3160.50*	*2*	*1580.25*	*55.02*	*< 0.0001*	
AB	*557.06*	*6*	*92.84*	*3.23*	*0.0180*	
Residual	689.33	24	28.72			
Lack of Fit	*0.000*	*0*				
Pure Error	*689.33*	*24*	*28.72*			
Cor Total	6532.00	35				

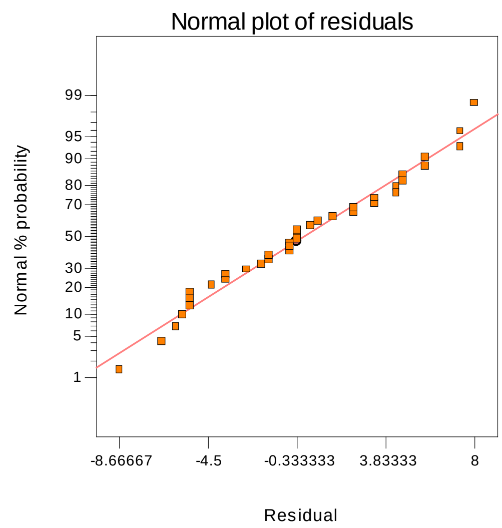
The Model F-value of 18.49 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

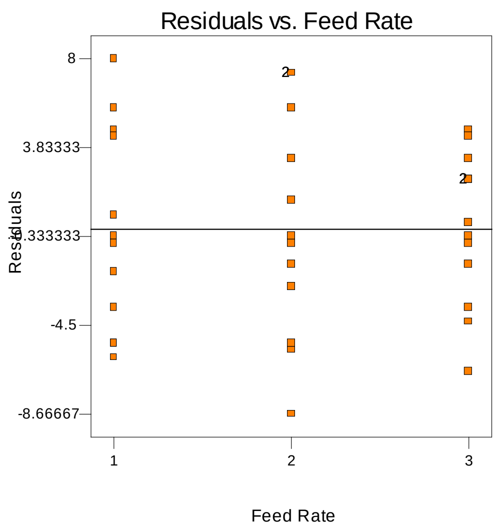
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, AB are significant model terms.

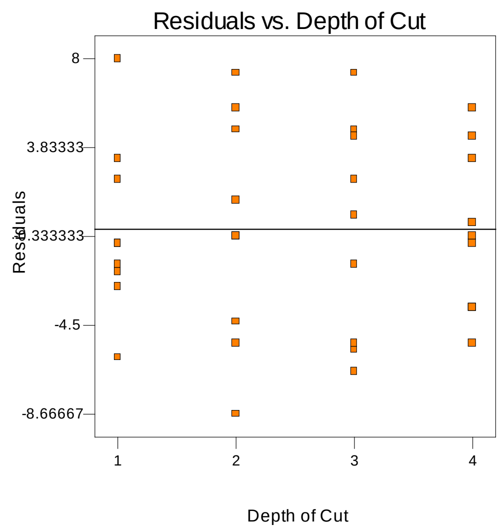
(b) Prepare appropriate residual plots and comment on the model's adequacy.

The residual plots shown indicate nothing unusual.



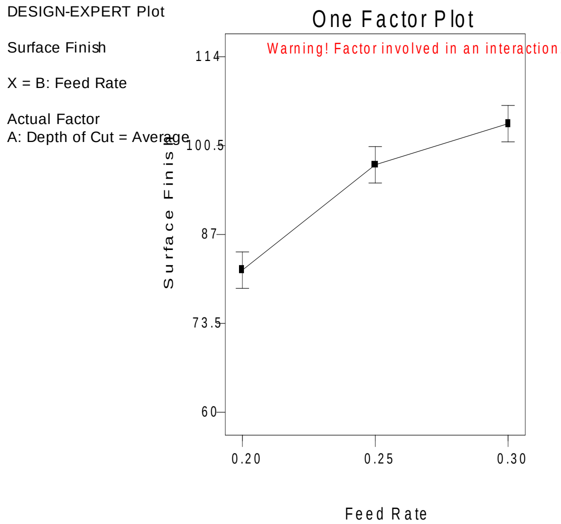






(c) Obtain point estimates of the mean surface finish at each feed rate.

Feed Rate	Average
0.20	81.58
0.25	97.58
0.30	103.83



(d) Find *P*-values for the tests in part (a).

The *P*-values are given in the computer output in part (a).

5.5. For the data in Problem 5.4, compute a 95 percent interval estimate of the mean difference in response for feed rates of 0.20 and 0.25 in/min.

We wish to find a confidence interval on , where μ_1 is the mean surface finish for 0.20 in/min and μ_2 is the mean surface finish for 0.25 in/min.

$$\bar{y}_{1..} - \bar{y}_{2..} - t_{\alpha/2,ab(n-1)}\sqrt{\frac{2MS_E}{n}} \le \mu_1 - \mu_2 \le \bar{y}_{1..} - \bar{y}_{2..} + t_{\alpha/2,ab(n-1))}\sqrt{\frac{2MS_E}{n}}$$

$$(81.5833 - 97.5833) \pm (2.064)\sqrt{\frac{2(28.7222)}{3}} = -16 \pm 9.032$$

Therefore, the 95% confidence interval for $\mu_1 - \mu_2$ is -16.000 ± 9.032.

5.6. An article in *Industrial Quality Control* (1956, pp. 5-8) describes an experiment to investigate the effect of the type of glass and the type of phosphor on the brightness of a television tube. The response variable is the current necessary (in microamps) to obtain a specified brightness level. The data are as follows:

Glass	Phosphor Type		
Type	1	2	3
	280	300	290
1	290	310	285
	285	295	290
	230	260	220
2	235	240	225
	240	235	230

(a) Is there any indication that either factor influences brightness? Use $\alpha = 0.05$.

Both factors, phosphor type (A) and Glass type (B) influence brightness.

Design Expert Output

Response: Current in microamps
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	15516.67	5	3103.33	58.80	< 0.0001	significant
A	*933.33*	*2*	*466.67*	*8.84*	*0.0044*	
B	*14450.00*	*1*	*14450.00*	*273.79*	*< 0.0001*	
AB	*133.33*	*2*	*66.67*	*1.26*	*0.3178*	
Residual	633.33	12	52.78			
Lack of Fit	*0.000*	*0*				
Pure Error	*633.33*	*12*	*52.78*			
Cor Total	16150.00	17				

The Model F-value of 58.80 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

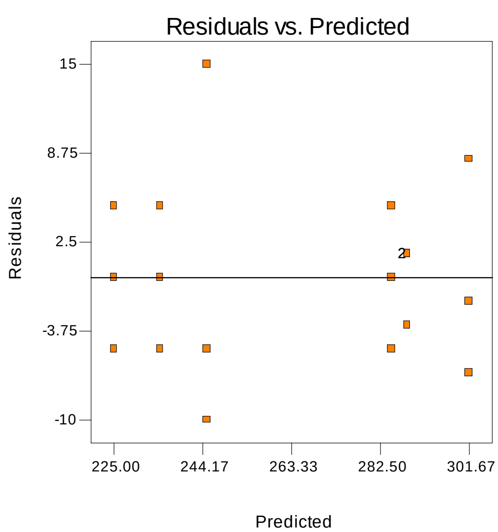
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

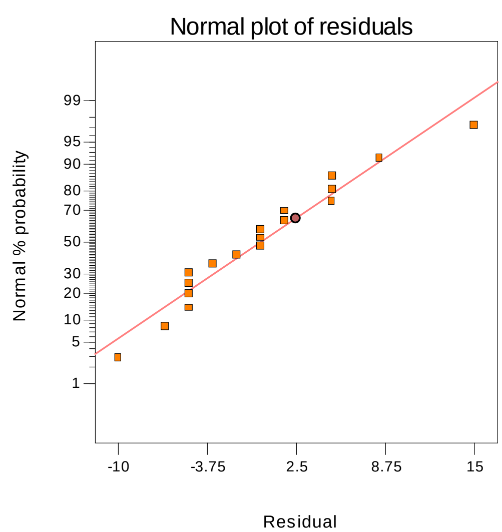
(b) Do the two factors interact? Use $\alpha = 0.05$.

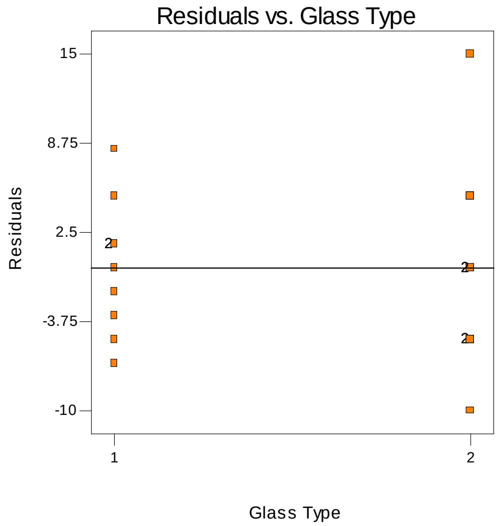
There is no interaction effect.

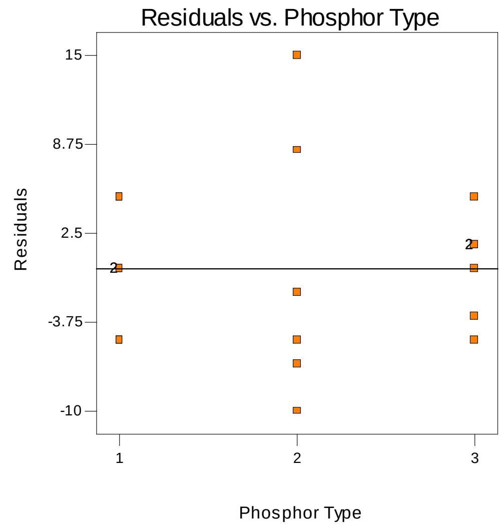
(c) Analyze the residuals from this experiment.

The residual plot of residuals versus phosphor content indicates a very slight inequality of variance. It is not serious enough to be of concern, however.









5.7. Johnson and Leone (*Statistics and Experimental Design in Engineering and the Physical Sciences*, Wiley 1977) describe an experiment to investigate the warping of copper plates. The two factors studied were the temperature and the copper content of the plates. The response variable was a measure of the amount of warping. The data were as follows:

	Copper Content (%)			
Temperature (°C)	40	60	80	100
50	17,20	16,21	24,22	28,27
75	12,9	18,13	17,12	27,31
100	16,12	18,21	25,23	30,23
125	21,17	23,21	23,22	29,31

(a) Is there any indication that either factor affects the amount of warping? Is there any interaction between the factors? Use $\alpha = 0.05$.

Both factors, copper content (*A*) and temperature (*B*) affect warping, the interaction does not.

Design Expert Output

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

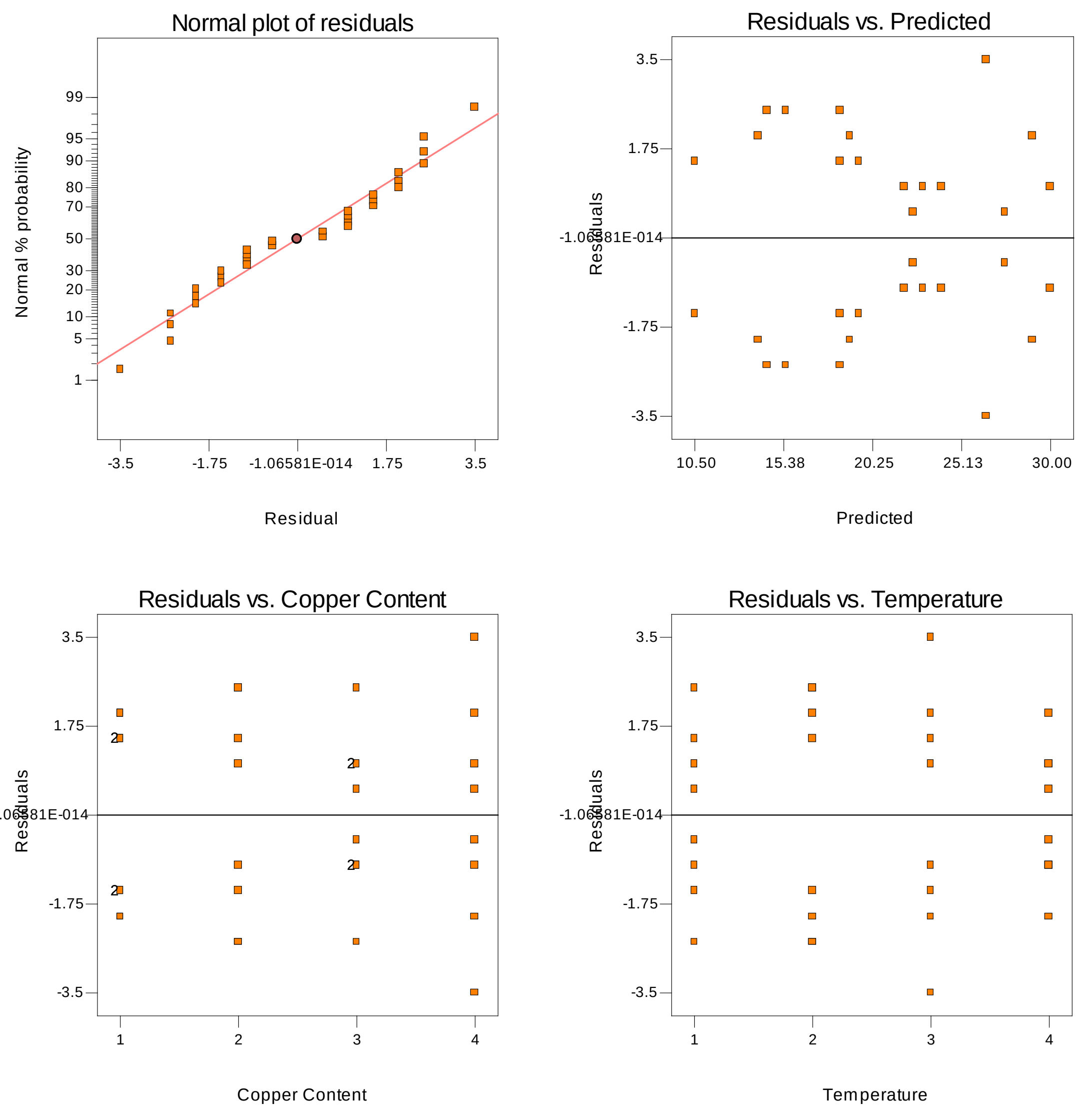
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	968.22	15	64.55	9.52	< 0.0001	significant
A	*698.34*	*3*	*232.78*	*34.33*	*< 0.0001*	
B	*156.09*	*3*	*52.03*	*7.67*	*0.0021*	
AB	*113.78*	*9*	*12.64*	*1.86*	*0.1327*	
Residual	108.50	16	6.78			
Lack of Fit	*0.000*	*0*				
Pure Error	*108.50*	*16*	*6.78*			
Cor Total	1076.72	31				

The Model F-value of 9.52 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.

(b) Analyze the residuals from this experiment.

There is nothing unusual about the residual plots.



(c) Plot the average warping at each level of copper content and compare them to an appropriately scaled *t* distribution. Describe the differences in the effects of the different levels of copper content on warping. If low warping is desirable, what level of copper content would you specify?

Design Expert Output

Factor	**Name**	**Level**	**Low Level**	**High Level**			
A	Copper Content	40	40	100			
B	Temperature	Average	50	125			
	Prediction	**SE Mean**	**95% CI low**	**95% CI high**	**SE Pred**	**95% PI low**	**95% PI high**
Warping	15.5	0.92	13.55	17.45	2.76	9.64	21.36
Factor	**Name**	**Level**	**Low Level**	**High Level**			
A	Copper Content	60	40	100			
B	Temperature	Average	50	125			

	Prediction	SE Mean	95% CI low	95% CI high	SE Pred	95% PI low	95% PI high
Warping	18.875	0.92	16.92	20.83	2.76	13.02	24.73

Factor	**Name**	**Level**	**Low Level**	**High Level**
A	**Copper Content**	**80**	**40**	**100**
B	**Temperature**	**Average**	**50**	**125**

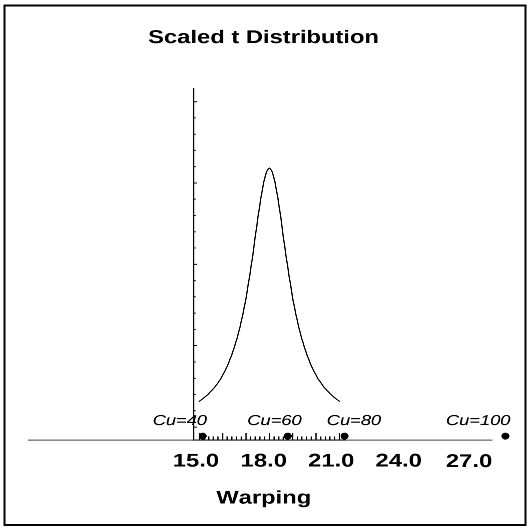
	Prediction	SE Mean	95% CI low	95% CI high	SE Pred	95% PI low	95% PI high
Warping	21	0.92	19.05	22.95	2.76	15.14	26.86

Factor	**Name**	**Level**	**Low Level**	**High Level**
A	**Copper Content**	**100**	**40**	**100**
B	**Temperature**	**Average**	**50**	**125**

	Prediction	SE Mean	95% CI low	95% CI high	SE Pred	95% PI low	95% PI high
Warping	28.25	0.92	26.30	30.20	2.76	22.39	34.11

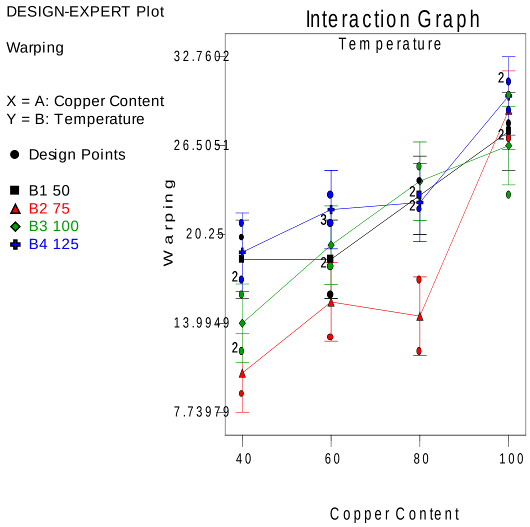
Use a copper content of 40 for the lowest warping.

$$S = \sqrt{\frac{MS_E}{b}} = \sqrt{\frac{6.78125}{8}} = 0.92$$



(d) Suppose that temperature cannot be easily controlled in the environment in which the copper plates are to be used. Does this change your answer for part (c)?

Use a copper of content of 40. This is the same as for part (c).



5.8. The factors that influence the breaking strength of a synthetic fiber are being studied. Four production machines and three operators are chosen and a factorial experiment is run using fiber from the same production batch. The results are as follows:

Operator	Machine			
	1	2	3	4
1	109	110	108	110
	110	115	109	108
2	110	110	111	114
	112	111	109	112
3	116	112	114	120
	114	115	119	117

(a) Analyze the data and draw conclusions. Use $\alpha = 0.05$.

Only the Operator (A) effect is significant.

Design Expert Output

Response:Stength
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	217.46	11	19.77	5.21	0.0041	significant
A	*160.33*	*2*	*80.17*	*21.14*	*0.0001*	
B	*12.46*	*3*	*4.15*	*1.10*	*0.3888*	
AB	*44.67*	*6*	*7.44*	*1.96*	*0.1507*	
Residual	45.50	12	3.79			
Lack of Fit	*0.000*	*0*				
Pure Error	*45.50*	*12*	*3.79*			
Cor Total	262.96	23				

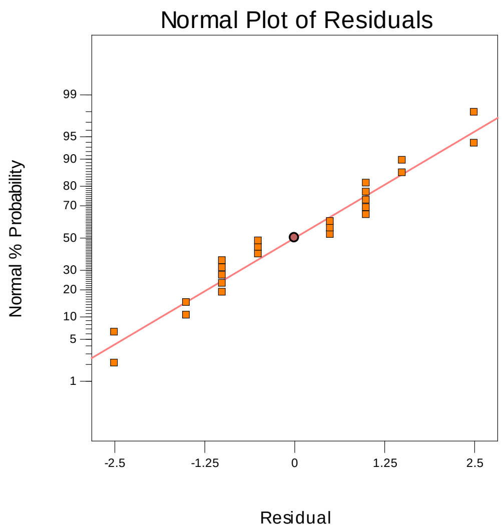
The Model F-value of 5.21 implies the model is significant.
There is only a 0.41% chance that a "Model F-Value" this large could occur due to noise.

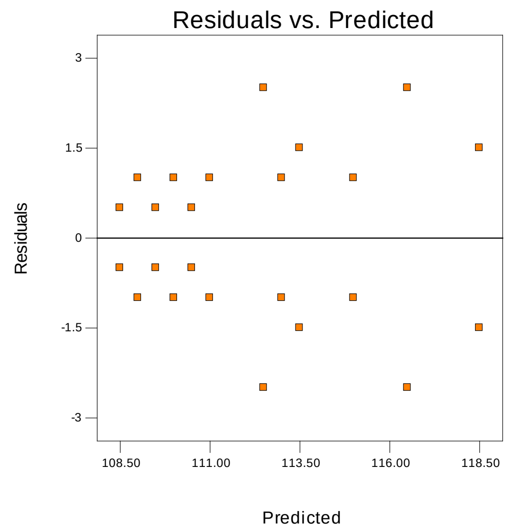
Values of "Prob > F" less than 0.0500 indicate model terms aresignificant.

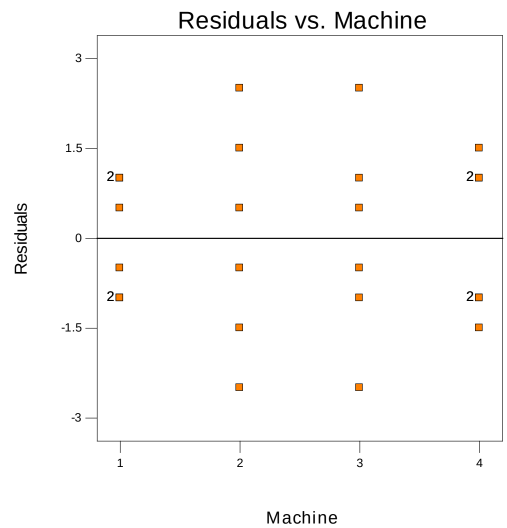
In this case A are significant model terms.

(b) Prepare appropriate residual plots and comment on the model's adequacy.

The residual plot of residuals versus predicted shows that variance increases very slightly with strength. There is no indication of a severe problem.







5.9. A mechanical engineer is studying the thrust force developed by a drill press. He suspects that the drilling speed and the feed rate of the material are the most important factors. He selects four feed rates and uses a high and low drill speed chosen to represent the extreme operating conditions. He obtains the following results. Analyze the data and draw conclusions. Use $\alpha = 0.05$.

(A)		Feed	Rate (B)	
Drill Speed	0.015	0.030	0.045	0.060
125	2.70	2.45	2.60	2.75

	2.78	2.49	2.72	2.86
200	2.83	2.85	2.86	2.94
	2.86	2.80	2.87	2.88

Design Expert Output

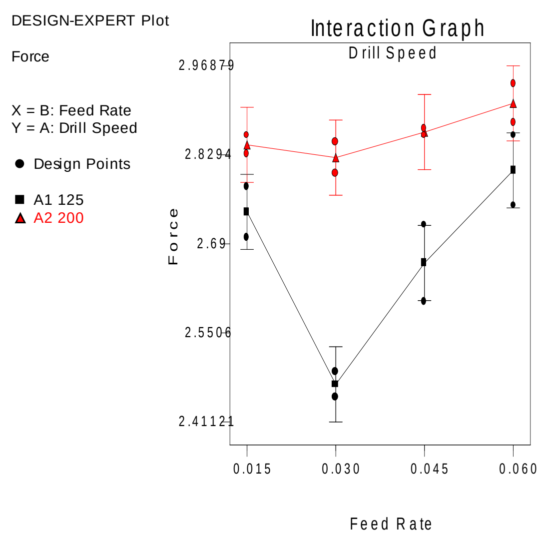
Response: Force
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	0.28	7	0.040	15.53	0.0005	significant
A	*0.15*	*1*	*0.15*	*57.01*	*< 0.0001*	
B	*0.092*	*3*	*0.031*	*11.86*	*0.0026*	
AB	*0.042*	*3*	*0.014*	*5.37*	*0.0256*	
Residual	0.021	8	2.600E-003			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.021*	*8*	*2.600E-003*			
Cor Total	0.30	15				

The Model F-value of 15.53 implies the model is significant.
There is only a 0.05% 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.

The factors speed and feed rate, as well as the interaction is important.



The standard analysis of variance treats all design factors as if they were qualitative. In this case, both factors are quantitative, so some further analysis can be performed. In Section 5.5, we show how response curves and surfaces can be fit to the data from a factorial experiment with at least one quantative factor. Since both factors in this problem are quantitative, we can fit polynomial effects of both speed and feed rate, exactly as in Example 5.5 in the text. The Design-Expert output with only the significant terms retained, including the response surface plots, now follows.

Design Expert Output

Response: Force
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	0.23	3	0.075	11.49	0.0008	significant
A	*0.15*	*1*	*0.15*	*22.70*	*0.0005*	
B	*0.019*	*1*	*0.019*	*2.94*	*0.1119*	

B2	*0.058*	*1*	*0.058*	*8.82*	*0.0117*	
Residual	0.078	12	6.530E-003			
Lack of Fit	*0.058*	*4*	*0.014*	*5.53*	*0.0196*	*significant*
Pure Error	*0.021*	*8*	*2.600E-003*			
Cor Total	0.30	15				

The Model F-value of 11.49 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^2 are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy),
model reduction may improve your model.

Std. Dev.	0.081	R-Squared	0.7417
Mean	2.77	Adj R-Squared	0.6772
C.V.	2.92	Pred R-Squared	0.5517
PRESS	0.14	Adeq Precision	9.269

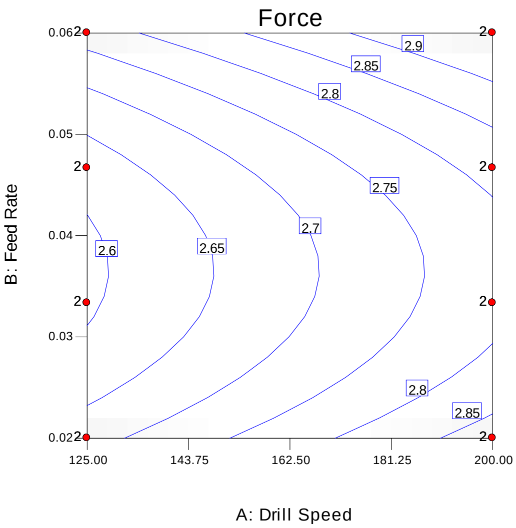
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	2.69	1	0.032	2.62	2.76	
A-Drill Speed	0.096	1	0.020	0.052	0.14	1.00
B-Feed Rate	0.047	1	0.027	-0.013	0.11	1.00
B2	0.13	1	0.045	0.036	0.23	1.00

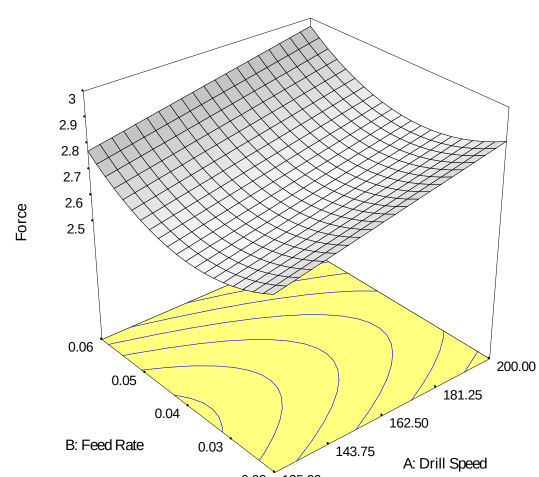
Final Equation in Terms of Coded Factors:

Force	=
+2.69	
+0.096	* A
+0.047	* B
+0.13	* B^2

Final Equation in Terms of Actual Factors:

Force	=
+2.48917	
+3.06667E-003	* Drill Speed
-15.76667	* Feed Rate
+266.66667	* Feed $Rate^2$





5.10. An experiment is conducted to study the influence of operating temperature and three types of face-plate glass in the light output of an oscilloscope tube. The following data are collected:

	Temperature		
Glass Type	100	125	150
1	580 568 570	1090 1087 1085	1392 1380 1386
2	550 530 579	1070 1035 1000	1328 1312 1299
3	546 575 599	1045 1053 1066	867 904 889

(a) Use $\alpha = 0.05$ in the analysis. Is there a significant interaction effect? Does glass type or temperature affect the response? What conclusions can you draw?

Design Expert Output

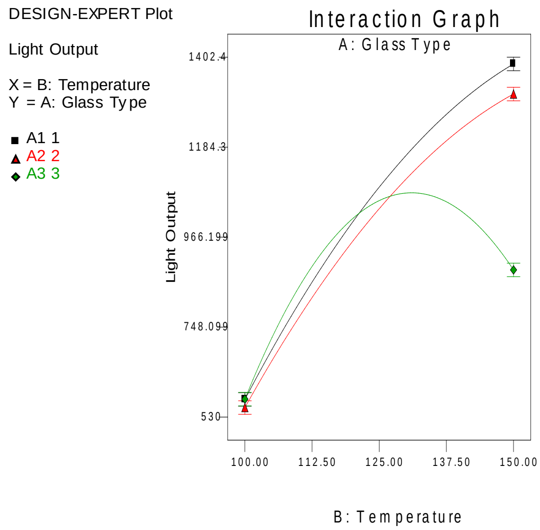
Response: Light Output
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2.412E+006	8	3.015E+005	824.77	< 0.0001	significant
A	*1.509E+005*	*2*	*75432.26*	*206.37*	*< 0.0001*	
B	*1.780E+006*	*1*	*1.780E+006*	*4869.13*	*< 0.0001*	
B2	*1.906E+005*	*1*	*1.906E+005*	*521.39*	*< 0.0001*	
AB	*2.262E+005*	*2*	*1.131E+005*	*309.39*	*< 0.0001*	
AB2	*64373.93*	*2*	*32186.96*	*88.06*	*< 0.0001*	
Pure Error	*6579.33*	*18*	*365.52*			
Cor Total	2.418E+006	26				

The Model F-value of 824.77 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	19.12	R-Squared	0.9973
Mean	940.19	Adj R-Squared	0.9961
C.V.	2.03	Pred R-Squared	0.9939
PRESS	14803.50	Adeq Precision	75.466

From the analysis of variance, both factors, Glass Type (*A*) and Temperature (*B*) are significant, as well as the interaction (*AB*). B^2 and AB^2 interation terms are also significant. The interaction and pure quadratic tems can be clearly seen in the plot shown below. For glass types 1 and 2 the temperature is fairly linear, for glass type 3, there is a quadratic effect.



(b) Fit an appropriate model relating light output to glass type and temperature.

The model, both coded and uncoded are shown in the *Design Expert* output below.

Design Expert Output

Final Equation in Terms of Coded Factors:

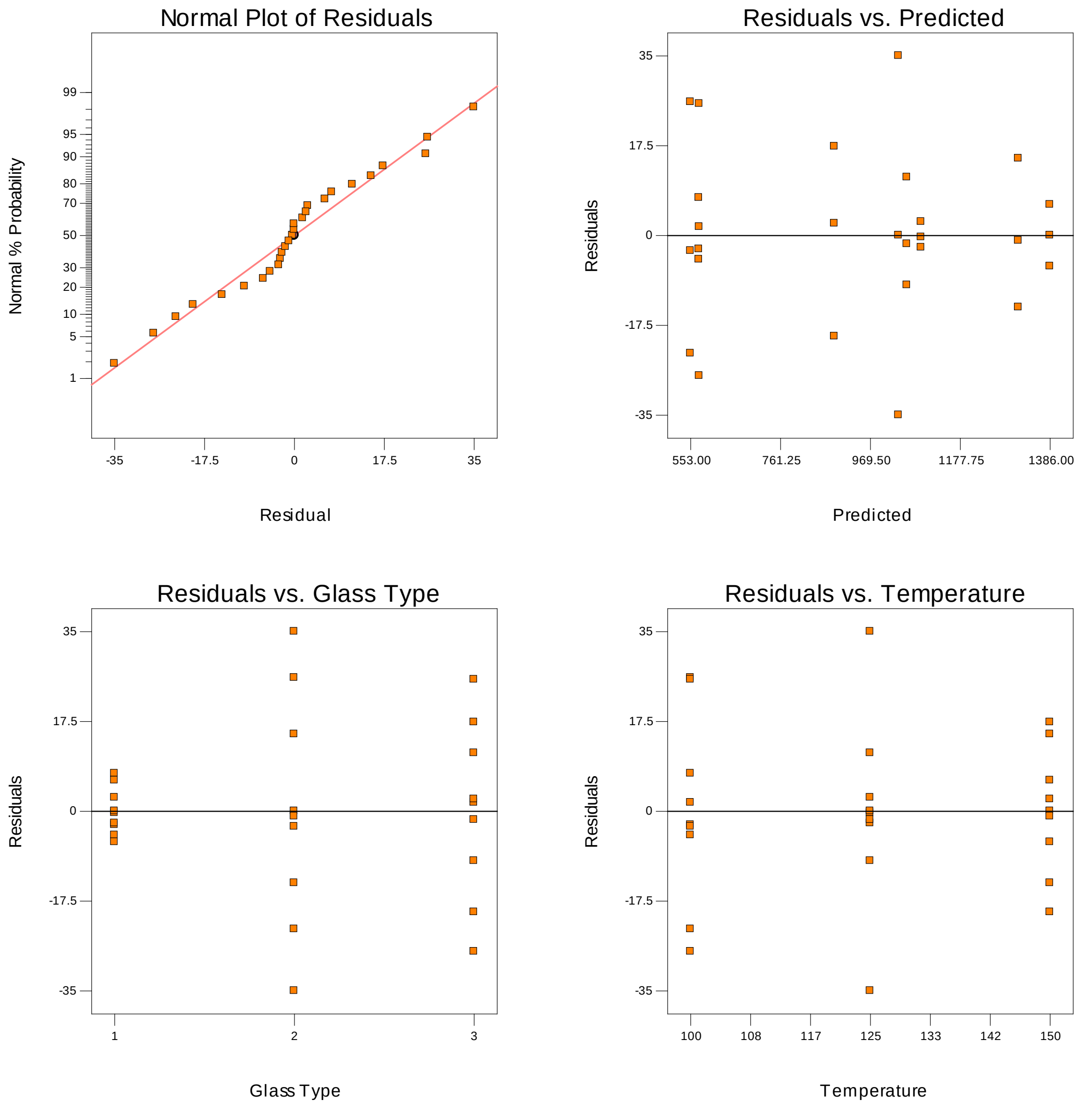
Light Output	=
+1059.00	
+28.33	* A[1]
-24.00	* A[2]
+314.44	* B
-178.22	* B2
+92.22	* A[1]B
+65.56	* A[2]B
+70.22	* A[1]B2
+76.22	* A[2]B2

Final Equation in Terms of Actual Factors:

Glass Type	1
Light Output	=
-3646.00000	
+59.46667	* Temperature
-0.17280	* Temperature2
Glass Type	2
Light Output	=
-3415.00000	
+56.00000	* Temperature
-0.16320	* Temperature2
Glass Type	3
Light Output	=
-7845.33333	
+136.13333	* Temperature
-0.51947	* Temperature2

(c) Analyze the residuals from this experiment. Comment on the adequacy of the models you have considered.

The only concern from the residuals below is the inequality of variance observed in the residuals versus glass type plot shown below.



5.11. Consider the data in Problem 5.3. Fit an appropriate model to the response data. Use this model to provide guidance concerning operating conditions for the process.

See the alternative analysis shown in Problem 5.3 part (c).

5.12. Use Tukey's test to determine which levels of the pressure factor are significantly different for the data in Problem 5.3.

Because the *AB* interaction is not significant, the sum of squares for the interaction is included as lack of fit in the residual error sum of squares for a SSE of 0.23 and degrees of freedom of 13. The sample size is also assumed to be increased from 2 to 6.

The three pressure averages, arranged in ascending order are

$$\bar{y}_{.3.} = 90.18 \qquad \bar{y}_{.1.} = 90.37 \qquad \bar{y}_{.2.} = 90.68$$

and

$$T_{0.05} = q_{0.05}(3,9)\sqrt{\frac{MS_E}{n}} = 3.95\sqrt{\frac{0.018}{6}} = 0.22$$

Comparing the differences with $T_{0.05}$, we have

$$\begin{aligned}
\bar{y}_{.2.} - \bar{y}_{.3.} &= 0.50 > T_{0.05} = 0.22 \\
\bar{y}_{.2.} - \bar{y}_{.1.} &= 0.32 > T_{0.05} = 0.22 \\
\bar{y}_{.1.} - \bar{y}_{.3.} &= 0.18 < T_{0.05} = 0.22
\end{aligned}$$

Therefore, the difference in yield between a pressure of 215 and 230 psig is statistically significant as is the difference in yield between 215 and 200 psig. However, the difference in yield between 200 and 230 psig is not statistically significant.

5.13. An experiment was conducted to determine if either firing temperature or furnace position affects the baked density of a carbon anode. The data are shown below.

	Temperature (°C)		
Position	800	825	850
	570	1063	565
1	565	1080	510
	583	1043	590
	528	988	526
2	547	1026	538
	521	1004	532

Suppose we assume that no interaction exists. Write down the statistical model. Conduct the analysis of variance and test hypotheses on the main effects. What conclusions can be drawn? Comment on the model's adequacy.

The model for the two-factor, no interaction model is $y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk}$. Both factors, furnace position (A) and temperature (B) are significant. Other than the residual representing standard order 14 being marginally low, the residual plots show nothing unusual.

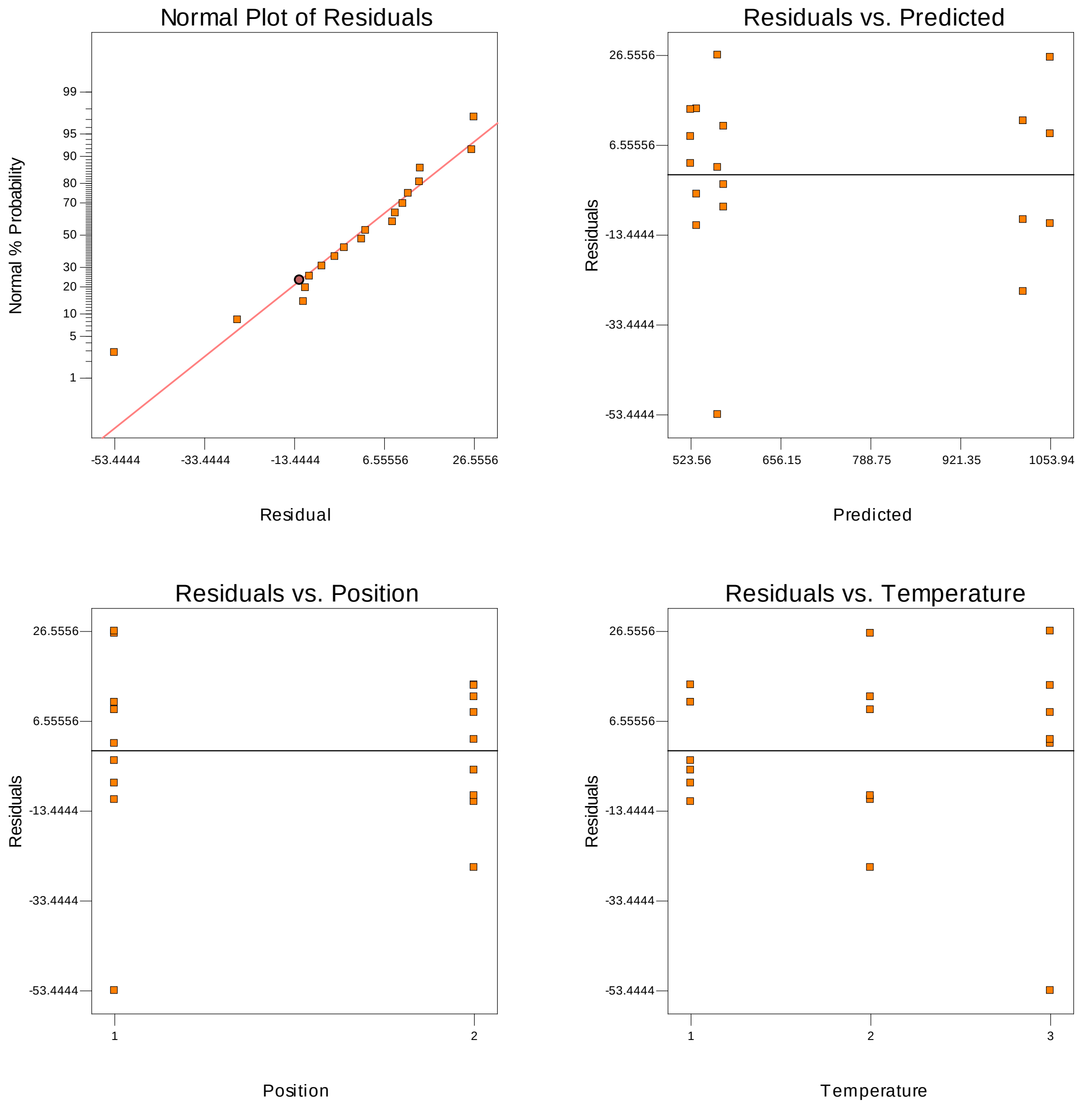
Design Expert Output

Response: Density
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	9.525E+005	3	3.175E+005	718.24	< 0.0001	significant
A	*7160.06*	*1*	*7160.06*	*16.20*	*0.0013*	
B	*9.453E+005*	*2*	*4.727E+005*	*1069.26*	*< 0.0001*	
Residual	6188.78	14	442.06			
Lack of Fit	*818.11*	*2*	*409.06*	*0.91*	*0.4271*	*not significant*
Pure Error	*5370.67*	*12*	*447.56*			
Cor Total	9.587E+005	17				

The Model F-value of 718.24 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.



5.14. Derive the expected mean squares for a two-factor analysis of variance with one observation per cell, assuming that both factors are fixed.

	Degrees of Freedom
$E(MS_A)=\sigma^2+b\sum_{i=1}^{a}\frac{\tau_i^2}{(a-1)}$	a-1
$E(MS_B)=\sigma^2+a\sum_{j=1}^{b}\frac{\beta_j^2}{(b-1)}$	b-1
$E(MS_{AB})=\sigma^2+\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(\tau\beta)_{ij}^2}{(a-1)(b-1)}$	$\frac{(a-1)(b-1)}{ab-1}$

5.15. Consider the following data from a two-factor factorial experiment. Analyze the data and draw conclusions. Perform a test for nonadditivity. Use $\alpha = 0.05$.

	Column	Factor		
Row Factor	1	2	3	4
1	36	39	36	32
2	18	20	22	20
3	30	37	33	34

Design Expert Output

Response: data
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	609.42	5	121.88	25.36	0.0006	significant
A	*580.50*	*2*	*290.25*	*60.40*	*0.0001*	
B	*28.92*	*3*	*9.64*	*2.01*	*0.2147*	
Residual	28.83	6	4.81			
Cor Total	638.25	11				

The Model F-value of 25.36 implies the model is significant. There is only a 0.06% chance that a "Model F-Value" this large could occur due to noise.

The row factor (A) is significant.

The test for nonadditivity is as follows:

$$SS_N=\frac{\left[\sum_{i=1}^{a}\sum_{j=1}^{b}y_{ij}y_{i.}y_{.j}-y_{..}\left(SS_A+SS_B+\frac{y_{..}^2}{ab}\right)\right]^2}{abSS_ASS_B}$$

$$SS_N=\frac{\left[4010014-(357)\left(580.50+28.91667+\frac{357^2}{(4)(3)}\right)\right]^2}{(4)(3)(580.50)(28.91667)}$$

$$SS_N=3.54051$$

$$SS_{Error}=SS_{Residual}-SS_N=28.8333-3.54051=25.29279$$

Source of	Sum of	Degrees of	Mean

Variation	Squares	Freedom	Square	F_0
Row	580.50	2	290.25	57.3780
Column	28.91667	3	9.63889	1.9054
Nonadditivity	3.54051	1	3.54051	0.6999
Error	25.29279	5	5.058558	
Total	638.25	11		

5.16. The shear strength of an adhesive is thought to be affected by the application pressure and temperature. A factorial experiment is performed in which both factors are assumed to be fixed. Analyze the data and draw conclusions. Perform a test for nonadditivity.

	Temperature (°F)		
Pressure (lb/in2)	250	260	270
120	9.60	11.28	9.00
130	9.69	10.10	9.57
140	8.43	11.01	9.03
150	9.98	10.44	9.80

Design Expert Output

Response: Strength
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	5.24	5	1.05	2.92	0.1124	not significant
A	*0.58*	*3*	*0.19*	*0.54*	*0.6727*	
B	*4.66*	*2*	*2.33*	*6.49*	*0.0316*	
Residual	2.15	6	0.36			
Cor Total	7.39	11				

The "Model F-value" of 2.92 implies the model is not significant relative to the noise.
There is a 11.24 % 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.

Temperature (*B*) is a significant factor.

$$SS_N = \frac{\left[\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij} y_{i.} y_{.j} - y_{..}\left(SS_A + SS_B + \frac{y_{..}^2}{ab}\right)\right]^2}{abSS_A SS_B}$$

$$SS_N = \frac{\left[137289.5797 - (117.93)\left(0.5806917 + 4.65765 + \frac{117.93^2}{(4)(3)}\right)\right]^2}{(4)(3)(0.5806917)(4.65765)}$$

$$SS_N = 0.48948$$

$$SS_{Error} = SS_{\mathrm{Re}sidual} - SS_N = 2.1538833 - 0.48948 = 1.66440$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Row	0.5806917	3	0.1935639	0.5815
Column	4.65765	2	2.328825	6.9960
Nonadditivity	0.48948	1	0.48948	1.4704
Error	1.6644	5	0.33288	
Total	7.392225	11		

5.17. Consider the three-factor model

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\beta\gamma)_{jk} + \varepsilon_{ijk} \quad \begin{cases} i = 1,2,...,a \\ j = 1,2,...,b \\ k = 1,2,...,c \end{cases}$$

Notice that there is only one replicate. Assuming the factors are fixed, write down the analysis of variance table, including the expected mean squares. What would you use as the "experimental error" in order to test hypotheses?

Source	Degrees of Freedom	Expected Mean Square
A	a-1	$\sigma^2 + bc\sum_{i=1}^{a}\frac{\tau_i^2}{(a-1)}$
B	b-1	$\sigma^2 + ac\sum_{j=1}^{b}\frac{\beta_j^2}{(b-1)}$
C	c-1	$\sigma^2 + ab\sum_{k=1}^{c}\frac{\gamma_k^2}{(c-1)}$
AB	(a-1)(b-1)	$\sigma^2 + c\sum_{i=1}^{a}\sum_{j=1}^{b}\frac{(\tau\beta)_{ij}^2}{(a-1)(b-1)}$
BC	(b-1)(c-1)	$\sigma^2 + a\sum_{j=1}^{b}\sum_{k=1}^{c}\frac{(\beta\gamma)_{jk}^2}{(b-1)(c-1)}$
Error (AC + ABC)	b(a-1)(c-1)	σ^2
Total	abc-1	

5.18. The percentage of hardwood concentration in raw pulp, the vat pressure, and the cooking time of the pulp are being investigated for their effects on the strength of paper. Three levels of hardwood concentration, three levels of pressure, and two cooking times are selected. A factorial experiment with two replicates is conducted, and the following data are obtained:

Percentage of Hardwood Concentration	Cooking Time 3.0 Hours			Cooking Time 4.0 Hours		
	Pressure			Pressure		
	400	500	650	400	500	650
2	196.6	197.7	199.8	198.4	199.6	200.6
	196.0	196.0	199.4	198.6	200.4	200.9
4	198.5	196.0	198.4	197.5	198.7	199.6
	197.2	196.9	197.6	198.1	198.0	199.0
8	197.5	195.6	197.4	197.6	197.0	198.5
	196.6	196.2	198.1	198.4	197.8	199.8

Percentage of Hardwood Concentration	Block	Cooking	Time 3.0 Pressure	Hours	Cooking	Time 4.0 Pressure	Hours
		400	500	650	400	500	650
2	1	196.6	197.7	199.8	198.4	199.6	200.6
	2	196.0	196.0	199.4	198.6	200.4	200.9
4	1	198.5	196.0	198.4	197.5	198.7	199.6
	2	197.2	196.9	197.6	198.1	198.0	199.0
8	1	197.5	195.6	197.4	197.6	197.0	198.5
	2	196.6	196.2	198.1	198.4	197.8	199.8

(a) Analyze the data and draw conclusions. Use $\alpha = 0.05$.

Design Expert Output

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

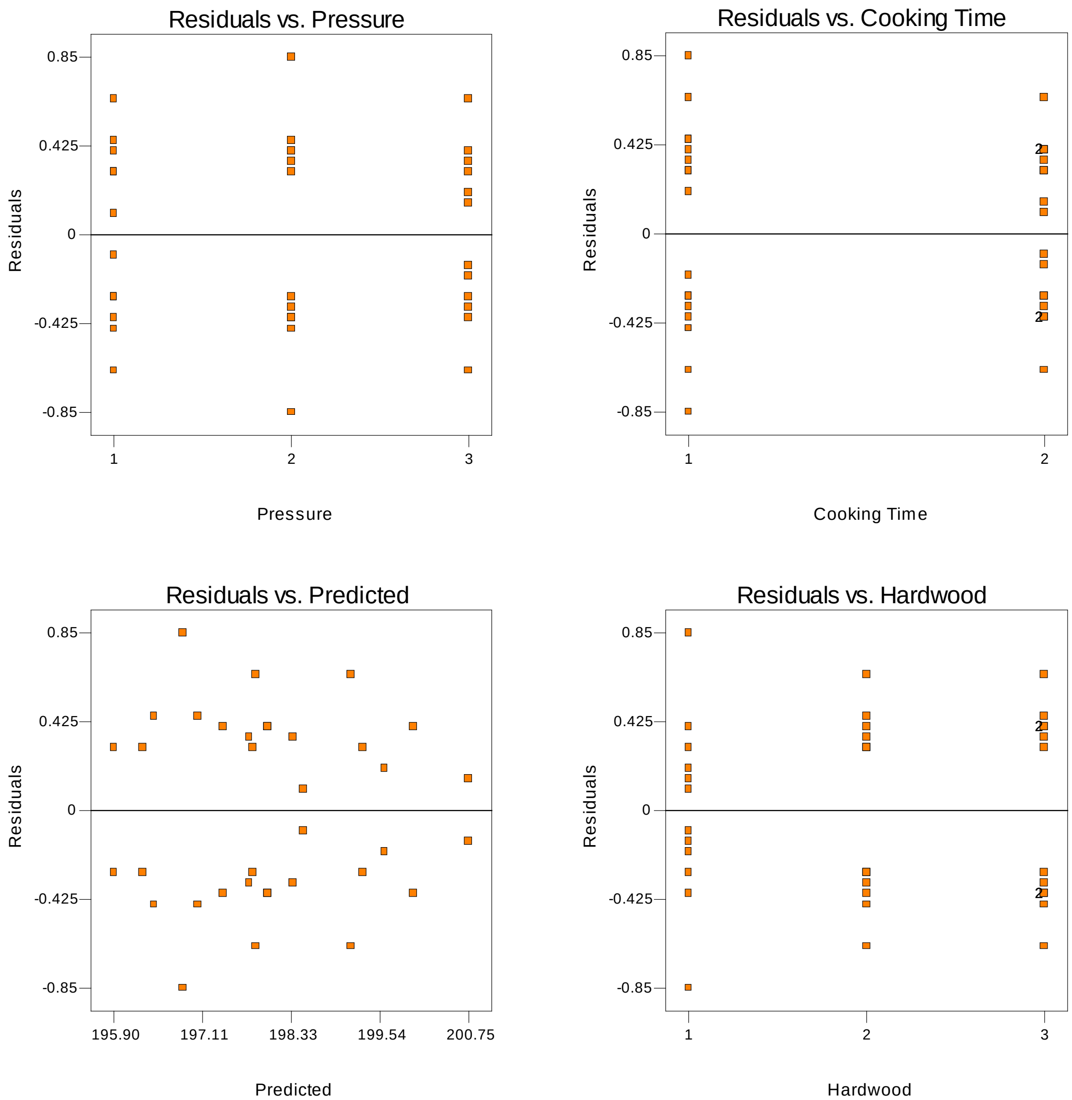
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	59.73	17	3.51	9.61	< 0.0001	significant
A	*7.76*	*2*	*3.88*	*10.62*	*0.0009*	
B	*20.25*	*1*	*20.25*	*55.40*	*< 0.0001*	
C	*19.37*	*2*	*9.69*	*26.50*	*< 0.0001*	
AB	*2.08*	*2*	*1.04*	*2.85*	*0.0843*	
AC	*6.09*	*4*	*1.52*	*4.17*	*0.0146*	
BC	*2.19*	*2*	*1.10*	*3.00*	*0.0750*	
ABC	*1.97*	*4*	*0.49*	*1.35*	*0.2903*	
Residual	6.58	18	0.37			
Lack of Fit	*0.000*	*0*				
Pure Error	*6.58*	*18*	*0.37*			
Cor Total	66.31	35				

The Model F-value of 9.61 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, AC are significant model terms.

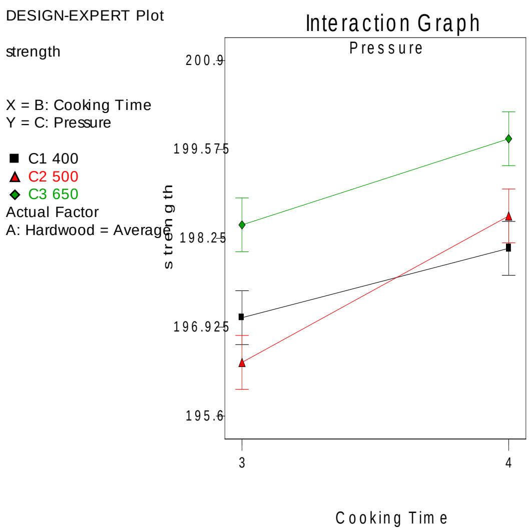
All three main effects, concentration (*A*), pressure (*C*) and time (*B*), as well as the concentration x pressure interaction (*AC*) are significant at the 5% level. The concentration x time (*AB*) and pressure x time interactions (*BC*) are significant at the 10% level.

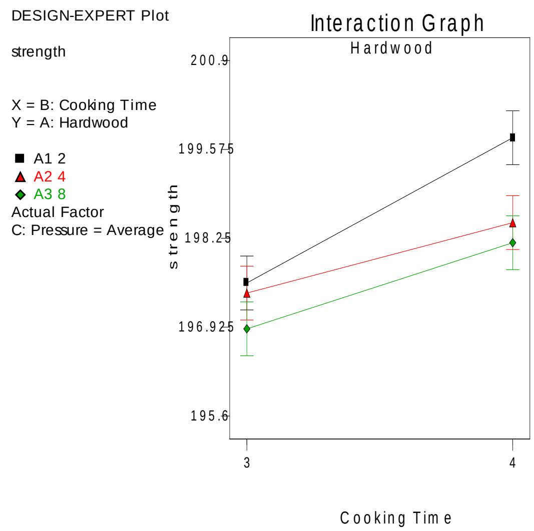
(b) Prepare appropriate residual plots and comment on the model's adequacy.

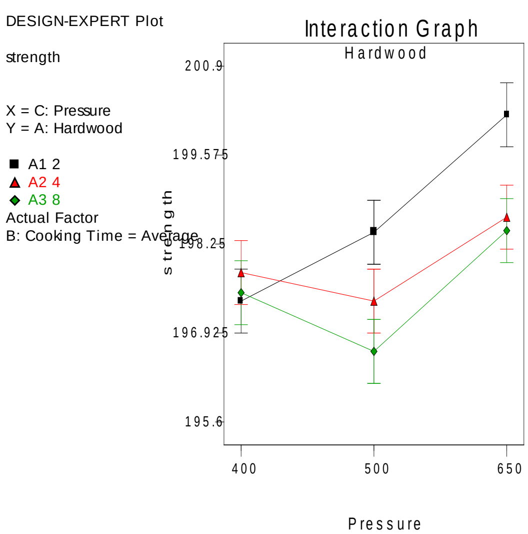


There is nothing unusual about the residual plots.

(c) Under what set of conditions would you run the process? Why?







For the highest strength, run the process with the percentage of hardwood at 2, the pressure at 650, and the time at 4 hours.

The standard analysis of variance treats all design factors as if they were qualitative. In this case, all three factors are quantitative, so some further analysis can be performed. In Section 5.5, we show how response curves and surfaces can be fit to the data from a factorial experiment with at least one quantative factor. Since the factors in this problem are quantitative and two of them have three levels, we can fit a linear term for the two-level factor and linear and quadratic components for the three-level factors. The Minitab output, with the *ABC* interaction removed due to insignificance, now follows. Also included is the Design Expert output; however, if the student choses to use Design Expert, sequential sum of squares must be selected to assure that the sum of squares for the model equals the total of the sum of squares for each factor included in the model.

Minitab Output

```
General Linear Model: Strength versus

Factor                      Type Levels Values

Analysis of Variance for Strength, using Adjusted SS for Tests

Source                      DF     Seq SS     Adj SS     Adj MS       F      P
Hardwood                     1     6.9067     4.9992     4.9992   13.23  0.001
Time                         1    20.2500     1.3198     1.3198    3.49  0.074
Pressure                     1    15.5605     1.5014     1.5014    3.97  0.058
Hardwood*Hardwood            1     0.8571     2.7951     2.7951    7.40  0.012
Pressure*Pressure            1     3.8134     1.8232     1.8232    4.83  0.038
Hardwood*Time                1     0.7779     1.5779     1.5779    4.18  0.053
Hardwood*Pressure            1     2.1179     3.4564     3.4564    9.15  0.006
Time*Pressure                1     0.0190     2.1932     2.1932    5.81  0.024
Hardwood*Hardwood*Time       1     1.3038     1.3038     1.3038    3.45  0.076
Hardwood*Hardwood*
Pressure                     1     2.1885     2.1885     2.1885    5.79  0.025
Hardwood*Pressure*
Pressure                     1     1.6489     1.6489     1.6489    4.36  0.048
Time*Pressure*Pressure       1     2.1760     2.1760     2.1760    5.76  0.025
Error                       23     8.6891     8.6891     0.3778
Total                       35    66.3089

Term                             Coef   SE Coef        T      P
Constant                       236.92     29.38     8.06  0.000
Hardwood                       10.728     2.949     3.64  0.001
Time                          -14.961     8.004    -1.87  0.074
Pressure                      -0.2257    0.1132    -1.99  0.058
Hardwood*Hardwood             -0.6529    0.2400    -2.72  0.012
Pressure*Pressure            0.000234  0.000107     2.20  0.038
Hardwood*Time                 -1.1750    0.5749    -2.04  0.053
Hardwood*Pressure           -0.020533  0.006788    -3.02  0.006
Time*Pressure                 0.07450   0.03092     2.41  0.024
Hardwood*Hardwood*Time        0.10278   0.05532     1.86  0.076
Hardwood*Hardwood*Pressure   0.000648  0.000269     2.41  0.025
Hardwood*Pressure*Pressure   0.000012  0.000006     2.09  0.048
Time*Pressure*Pressure      -0.000070  0.000029    -2.40  0.025

Unusual Observations for Strength

Obs  Strength       Fit      SE Fit  Residual   St Resid
  6   198.500   197.461       0.364     1.039      2.10R

R denotes an observation with a large standardized residual.
```

Design Expert Output

Response: Strength
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	57.62	12	4.80	12.71	< 0.0001	significant
A	*6.91*	*1*	*6.91*	*18.28*	*0.0003*	
B	*20.25*	*1*	*20.25*	*53.60*	*< 0.0001*	
C	*15.56*	*1*	*15.56*	*41.19*	*< 0.0001*	
A2	*0.86*	*1*	*0.86*	*2.27*	*0.1456*	
C2	*3.81*	*1*	*3.81*	*10.09*	*0.0042*	
AB	*0.78*	*1*	*0.78*	*2.06*	*0.1648*	
AC	*2.12*	*1*	*2.12*	*5.61*	*0.0267*	
BC	*0.019*	*1*	*0.019*	*0.050*	*0.8245*	
A2B	*1.30*	*1*	*1.30*	*3.45*	*0.0761*	
A2C	*2.19*	*1*	*2.19*	*5.79*	*0.0245*	
AC2	*1.65*	*1*	*1.65*	*4.36*	*0.0479*	
BC2	*2.18*	*1*	*2.18*	*5.76*	*0.0249*	
Residual	8.69	23	0.38			
Lack of Fit	*2.11*	*5*	*0.42*	*1.15*	*0.3691*	*not significant*
Pure Error	*6.58*	*18*	*0.37*			
Cor Total	66.31	35				

The Model F-value of 12.71 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^2, AC, A^2C, AC^2, BC^2 are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy),
model reduction may improve your model.

Std. Dev.	0.61	R-Squared	0.8690
Mean	198.06	Adj R-Squared	0.8006
C.V.	0.31	Pred R-Squared	0.6794
PRESS	21.26	Adeq Precision	15.040

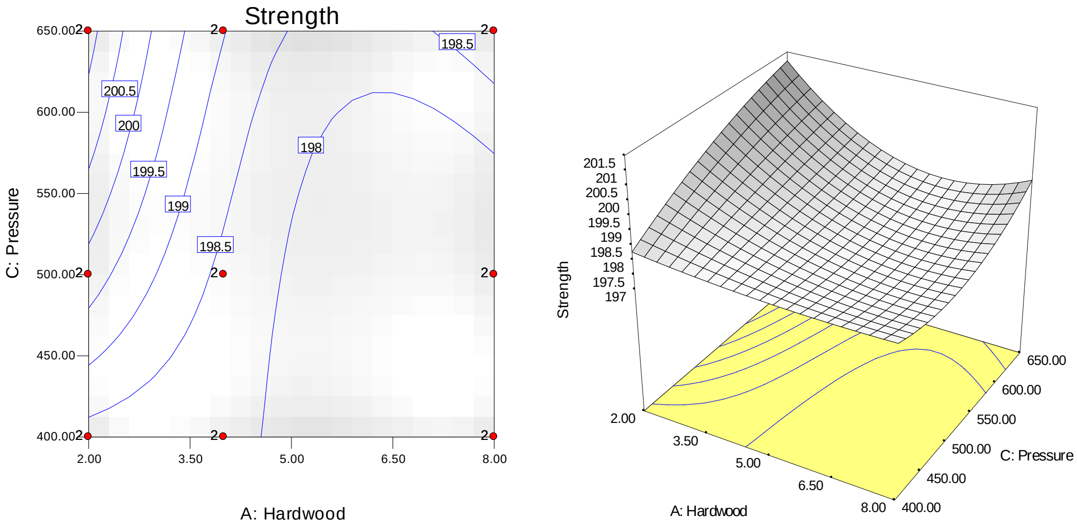
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	197.21	1	0.26	196.67	197.74	
A-Hardwood	-0.98	1	0.23	-1.45	-0.52	3.36
B-Cooking Time	0.78	1	0.26	0.24	1.31	6.35
C-Pressure	0.19	1	0.25	-0.33	0.71	4.04
A2	0.42	1	0.25	-0.093	0.94	1.04
C2	0.79	1	0.23	0.31	1.26	1.03
AB	-0.22	1	0.13	-0.48	0.039	1.06
AC	-0.46	1	0.15	-0.78	-0.14	1.08
BC	0.062	1	0.13	-0.20	0.32	1.02
A2B	0.46	1	0.25	-0.053	0.98	3.96
A2C	0.73	1	0.30	0.10	1.36	3.97
AC2	0.57	1	0.27	5.625E-003	1.14	3.32
BC2	-0.55	1	0.23	-1.02	-0.075	3.30

Final Equation in Terms of Coded Factors:

Strength	=
+197.21	
-0.98	* A
+0.78	* B
+0.19	* C
+0.42	* A2
+0.79	* C2
-0.22	* A * B
-0.46	* A * C
+0.062	* B * C
+0.46	* A2 * B
+0.73	* A2 * C
+0.57	* A * C2
-0.55	* B * C2

Final Equation in Terms of Actual Factors:

Strength	=
+236.91762	
+10.72773	* Hardwood
-14.96111	* Cooking Time
-0.22569	* Pressure
-0.65287	* Hardwood2
+2.34333E-004	* Pressure2
-1.17500	* Hardwood * Cooking Time
-0.020533	* Hardwood * Pressure
+0.074500	* Cooking Time * Pressure
+0.10278	* Hardwood2 * Cooking Time
+6.48026E-004	* Hardwood2 * Pressure
+1.22143E-005	* Hardwood * Pressure2
-7.00000E-005	* Cooking Time * Pressure2



Cooking Time: B = 4.00

5.19. The quality control department of a fabric finishing plant is studying the effect of several factors on the dyeing of cotton-synthetic cloth used to manufacture men's shirts. Three operators, three cycle times, and two temperatures were selected, and three small specimens of cloth were dyed under each set of conditions. The finished cloth was compared to a standard, and a numerical score was assigned. The results follow. Analyze the data and draw conclusions. Comment on the model's adequacy.

	Temperature					
	300°C			350°C		
	Operator			Operator		
Cycle Time	1	2	3	1	2	3
40	23	27	31	24	38	34
	24	28	32	23	36	36
	25	26	29	28	35	39
50	36	34	33	37	34	34
	35	38	34	39	38	36
	36	39	35	35	36	31
60	28	35	26	26	36	28
	24	35	27	29	37	26
	27	34	25	25	34	24

All three main effects, and the *AB*, *AC*, and *ABC* interactions are significant. There is nothing unusual about the residual plots.

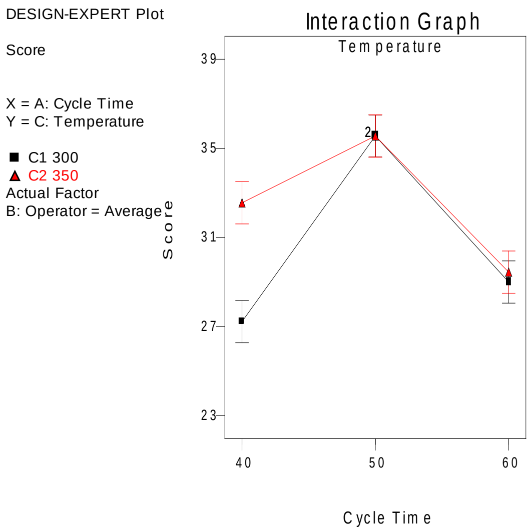
Design Expert Output

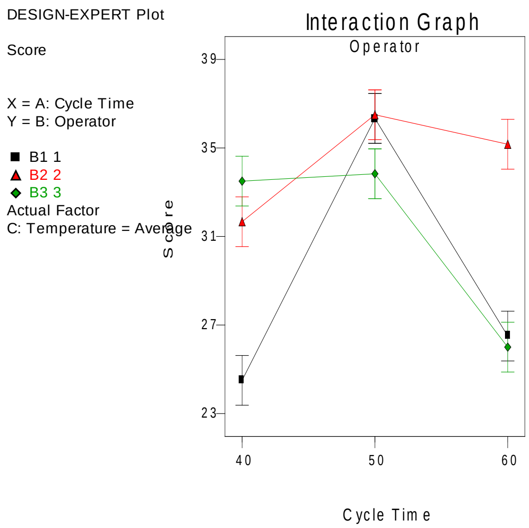
Response: Score
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

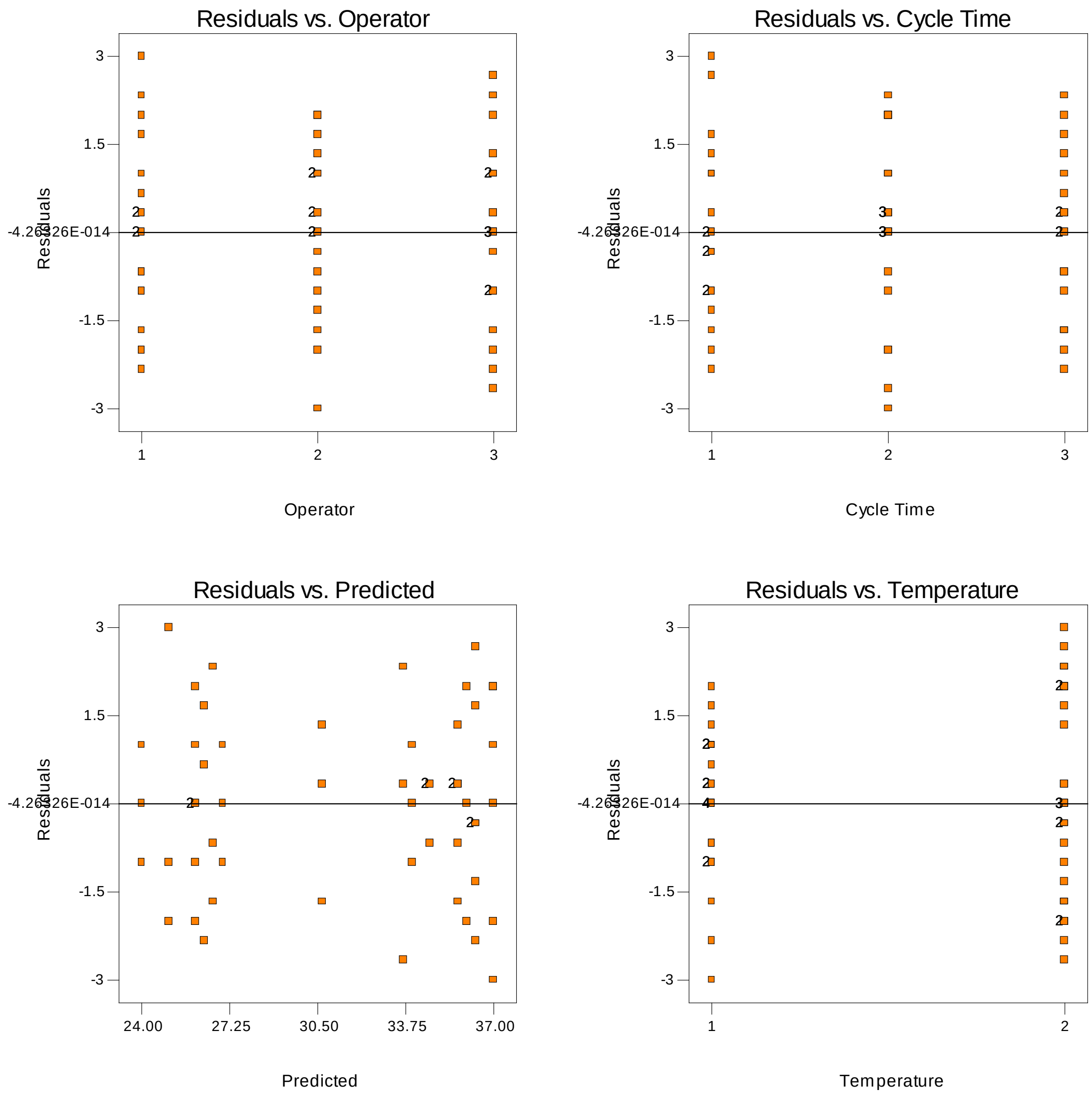
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1239.33	17	72.90	22.24	< 0.0001	significant
A	*436.00*	*2*	*218.00*	*66.51*	*< 0.0001*	
B	*261.33*	*2*	*130.67*	*39.86*	*< 0.0001*	
C	*50.07*	*1*	*50.07*	*15.28*	*0.0004*	
AB	*355.67*	*4*	*88.92*	*27.13*	*< 0.0001*	
AC	*78.81*	*2*	*39.41*	*12.02*	*0.0001*	
BC	*11.26*	*2*	*5.63*	*1.72*	*0.1939*	
ABC	*46.19*	*4*	*11.55*	*3.52*	*0.0159*	
Residual	118.00	36	3.28			
Lack of Fit	*0.000*	*0*				
Pure Error	*118.00*	*36*	*3.28*			
Cor Total	1357.33	53				

The Model F-value of 22.24 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, AC, ABC are significant model terms.







5.20. In Problem 5.3, suppose that we wish to reject the null hypothesis with a high probability if the difference in the true mean yield at any two pressures is as great as 0.5. If a reasonable prior estimate of the standard deviation of yield is 0.1, how many replicates should be run?

$$\Phi^2 = \frac{naD^2}{2b\sigma^2} = \frac{n(3)(0.5)^2}{2(3)(0.1)^2} = 12.5n$$

n	Φ^2	Φ	$\upsilon_1 = (b-1)$	$\upsilon_2 = ab(n-1)$	β
2	25	5	2	(3)(3)(1)	0.014

2 replications will be enough to detect the given difference.

5.21. The yield of a chemical process is being studied. The two factors of interest are temperature and pressure. Three levels of each factor are selected; however, only 9 runs can be made in one day. The experimenter runs a complete replicate of the design on each day. The data are shown in the following table. Analyze the data assuming that the days are blocks.

	Day 1 Pressure			Day 2 Pressure		
Temperature	250	260	270	250	260	270
Low	86.3	84.0	85.8	86.1	85.2	87.3
Medium	88.5	87.3	89.0	89.4	89.9	90.3
High	89.1	90.2	91.3	91.7	93.2	93.7

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**	
Block	13.01	1	13.01			
Model	109.81	8	13.73	25.84	< 0.0001	significant
A	*5.51*	*2*	*2.75*	*5.18*	*0.0360*	
B	*99.85*	*2*	*49.93*	*93.98*	*< 0.0001*	
AB	*4.45*	*4*	*1.11*	*2.10*	*0.1733*	
Residual	4.25	8	0.53			
Cor Total	127.07	17				

The Model F-value of 25.84 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 are significant model terms.

Both main effects, temperature and pressure, are significant.

5.22. Consider the data in Problem 5.7. Analyze the data, assuming that replicates are blocks.

Design Expert Output

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

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	11.28	1	11.28			
Model	968.22	15	64.55	9.96	< 0.0001	significant
A	*698.34*	*3*	*232.78*	*35.92*	*< 0.0001*	
B	*156.09*	*3*	*52.03*	*8.03*	*0.0020*	
AB	*113.78*	*9*	*12.64*	*1.95*	*0.1214*	
Residual	97.22	15	6.48			
Cor Total	1076.72	31				

The Model F-value of 9.96 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.

Both temperature and copper content are significant. This agrees with the analysis in Problem 5.7.

5.23. Consider the data in Problem 5.8. Analyze the data, assuming that replicates are blocks.

Design-Expert Output

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

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	2.04	1	2.04			
Model	217.46	11	19.77	5.00	0.0064	significant
A	*160.33*	*2*	*80.17*	*20.29*	*0.0002*	
B	*12.46*	*3*	*4.15*	*1.05*	*0.4087*	
AB	*44.67*	*6*	*7.44*	*1.88*	*0.1716*	
Residual	43.46	11	3.95			
Cor Total	262.96	23				

The Model F-value of 5.00 implies the model is significant. There is only
a 0.64% 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.

Only the operator factor (*A*) is significant. This agrees with the analysis in Problem 5.8.

5.24. An article in the *Journal of Testing and Evaluation* (Vol. 16, no.2, pp. 508-515) investigated the effects of cyclic loading and environmental conditions on fatigue crack growth at a constant 22 MPa stress for a particular material. The data from this experiment are shown below (the response is crack growth rate).

	Environment		
Frequency	Air	H_2O	Salt H_2O
10	2.29	2.06	1.90
	2.47	2.05	1.93
	2.48	2.23	1.75
	2.12	2.03	2.06
1	2.65	3.20	3.10
	2.68	3.18	3.24
	2.06	3.96	3.98
	2.38	3.64	3.24
0.1	2.24	11.00	9.96
	2.71	11.00	10.01
	2.81	9.06	9.36
	2.08	11.30	10.40

(a) Analyze the data from this experiment (use $\alpha = 0.05$).

Design Expert Output

Response: Crack 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	376.11	8	47.01	234.02	< 0.0001	significant
A	*209.89*	*2*	*104.95*	*522.40*	*< 0.0001*	
B	*64.25*	*2*	*32.13*	*159.92*	*< 0.0001*	
AB	*101.97*	*4*	*25.49*	*126.89*	*< 0.0001*	
Residual	5.42	27	0.20			
Lack of Fit	*0.000*	*0*				
Pure Error	*5.42*	*27*	*0.20*			
Cor Total	381.53	35				

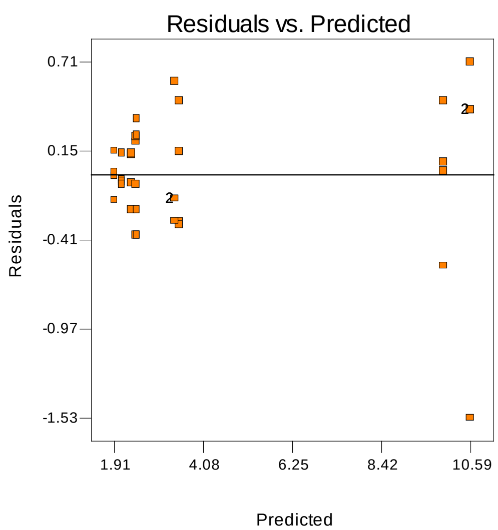
The Model F-value of 234.02 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

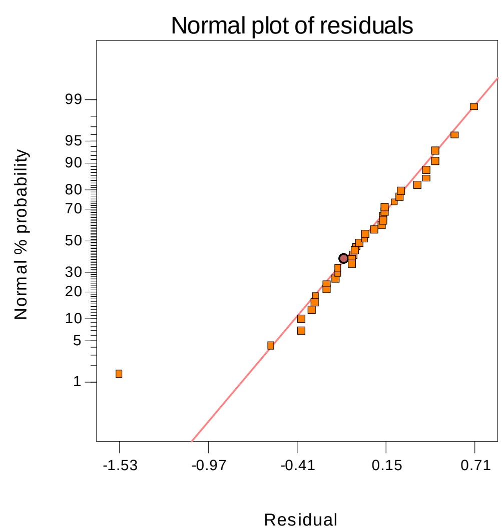
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, AB are significant model terms.

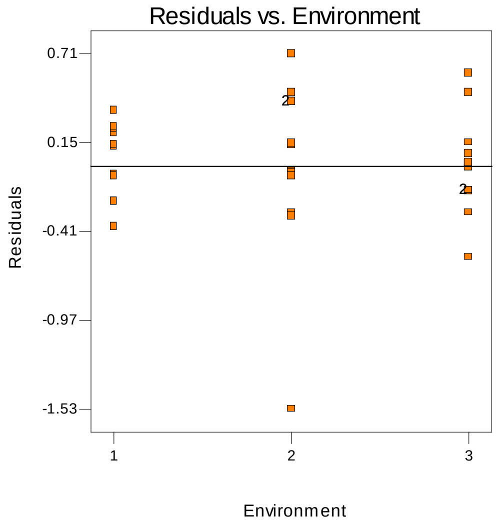
Both frequency and environment, as well as their interaction are significant.

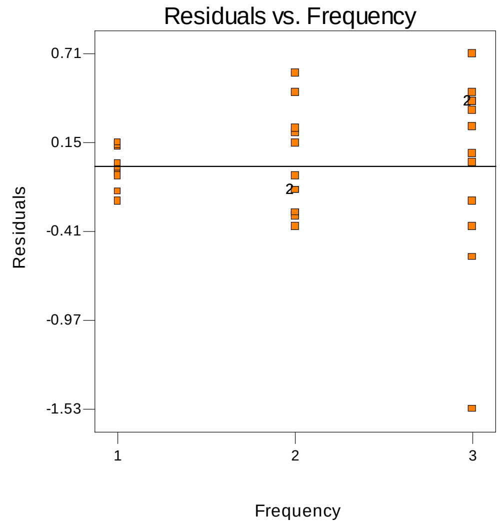
(b) Analyze the residuals.

The residual plots indicate that there may be some problem with inequality of variance. This is particularly noticable on the plot of residuals versus predicted response and the plot of residuals versus frequency.









(c) Repeat the analyses from parts (a) and (b) using ln(y) as the response. Comment on the results.

Design Expert Output

Response: Crack Growth **Transform: Natural log** **Constant: 0.000**

ANOVA for Selected Factorial Model

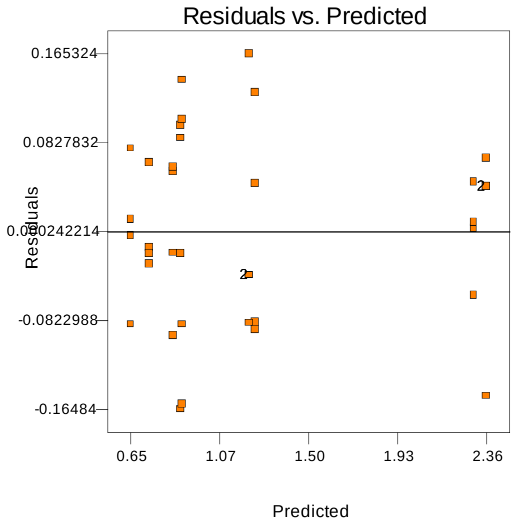
Analysis of variance table [Partial sum of squares]

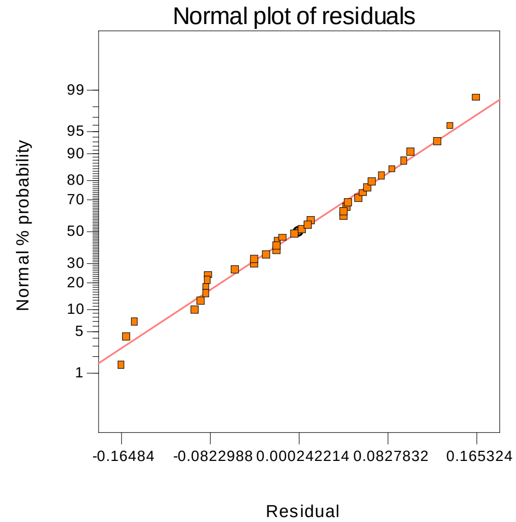
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	13.46	8	1.68	179.57	< 0.0001	significant
A	*7.57*	*2*	*3.79*	*404.09*	*< 0.0001*	
B	*2.36*	*2*	*1.18*	*125.85*	*< 0.0001*	
AB	*3.53*	*4*	*0.88*	*94.17*	*< 0.0001*	
Residual	0.25	27	9.367E-003			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.25*	*27*	*9.367E-003*			
Cor Total	13.71	35				

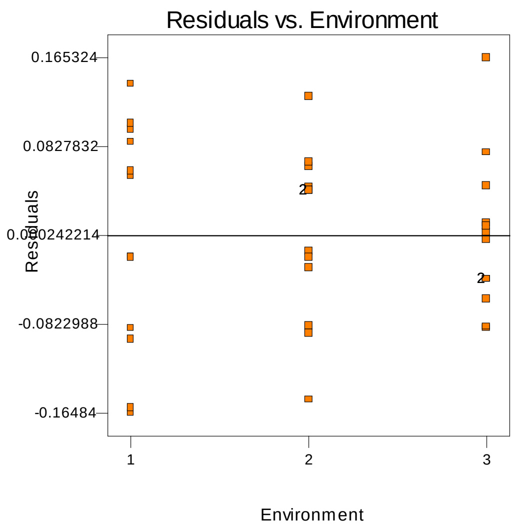
The Model F-value of 179.57 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

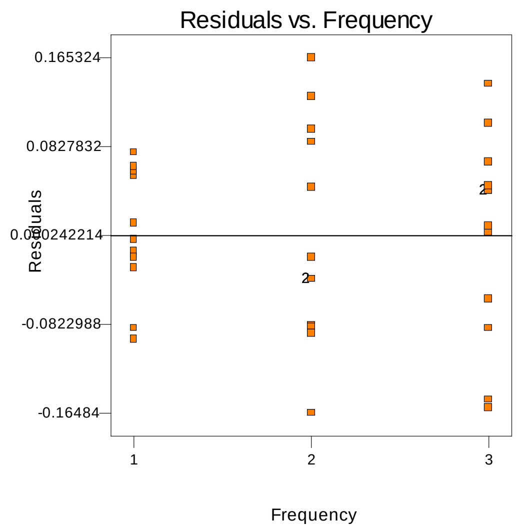
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, AB are significant model terms.

Both frequency and environment, as well as their interaction are significant. The residual plots based on the transformed data look better.









5.25. An article in the *IEEE Transactions on Electron Devices* (Nov. 1986, pp. 1754) describes a study on polysilicon doping. The experiment shown below is a variation of their study. The response variable is base current.

Polysilicon	Anneal Temperature (°C)		
Doping (ions)	900	950	1000
1×10^{20}	4.60	10.15	11.01
	4.40	10.20	10.58
2×10^{20}	3.20	9.38	10.81
	3.50	10.02	10.60

(a) Is there evidence (with $\alpha = 0.05$) indicating that either polysilicon doping level or anneal temperature affect base current?

Design Expert Output

Response: Base Current
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

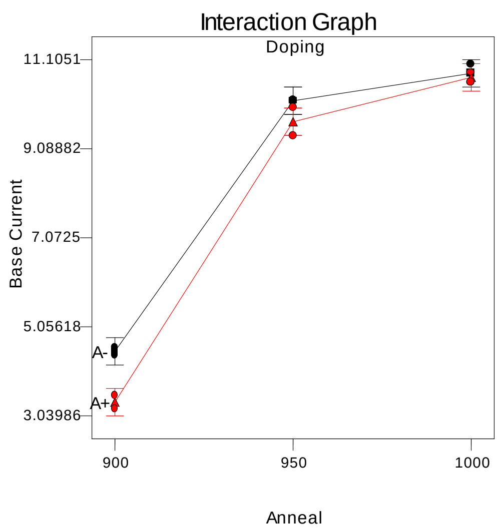
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	112.74	5	22.55	350.91	< 0.0001	significant
A	*0.98*	*1*	*0.98*	*15.26*	*0.0079*	
B	*111.19*	*2*	*55.59*	*865.16*	*< 0.0001*	
AB	*0.58*	*2*	*0.29*	*4.48*	*0.0645*	
Residual	0.39	6	0.064			
Lack of Fit	*0.000*	*0*				
Pure Error	*0.39*	*6*	*0.064*			
Cor Total	113.13	11				

The Model F-value of 350.91 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

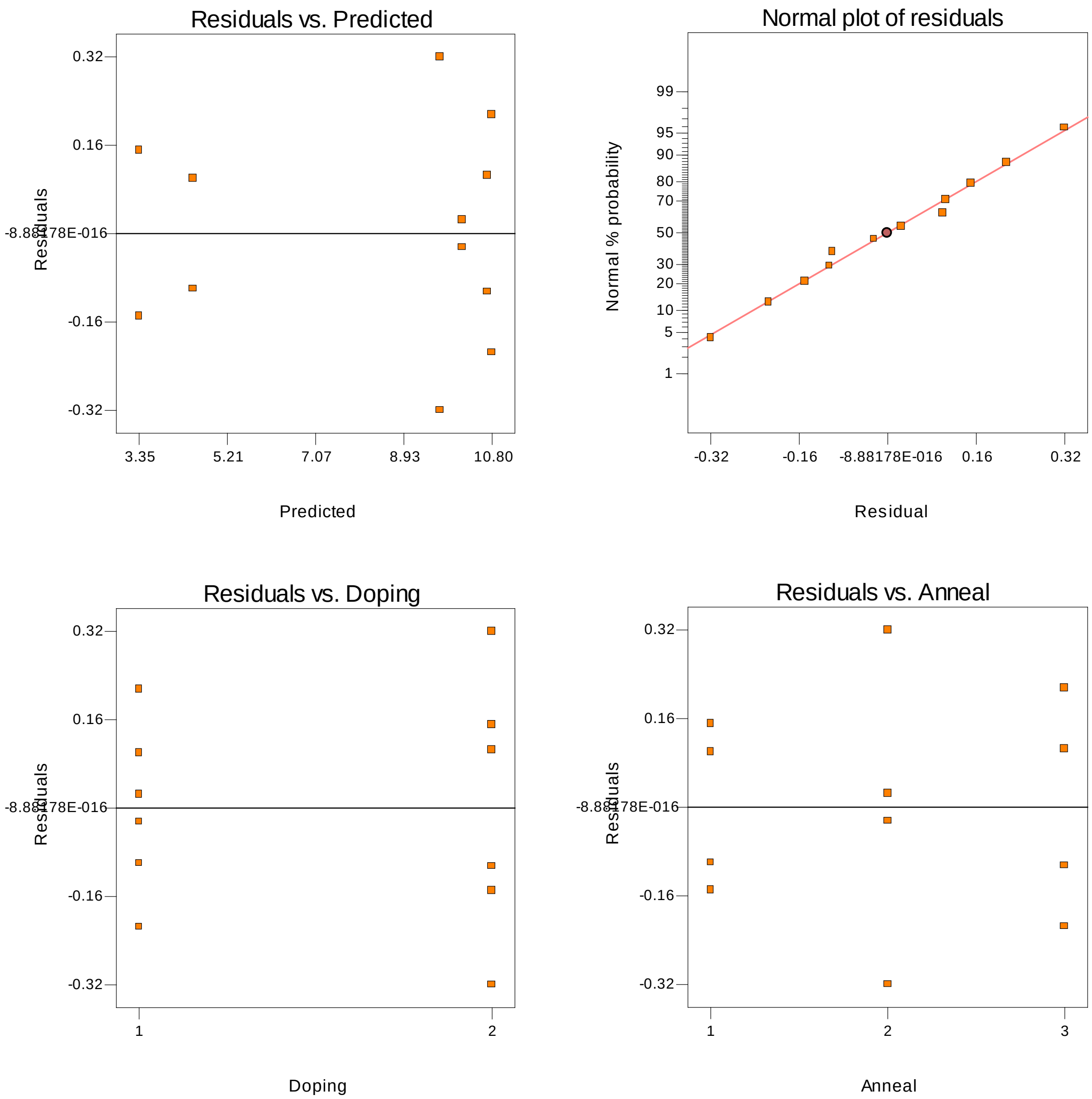
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

Both factors, doping and anneal are significant. Their interaction is significant at the 10% level.

(b) Prepare graphical displays to assist in interpretation of this experiment.



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



There is a funnel shape in the plot of residuals versus predicted, indicating some inequality of variance.

(d) Is the model $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$ supported by this experiment (x_1 = doping level, x_2 = temperature)? Estimate the parameters in this model and plot the response surface.

Design Expert Output

Response: **Base Current**
ANOVA for Response Surface Reduced Quadratic Model
Analysis of variance table [Partial sum of squares]

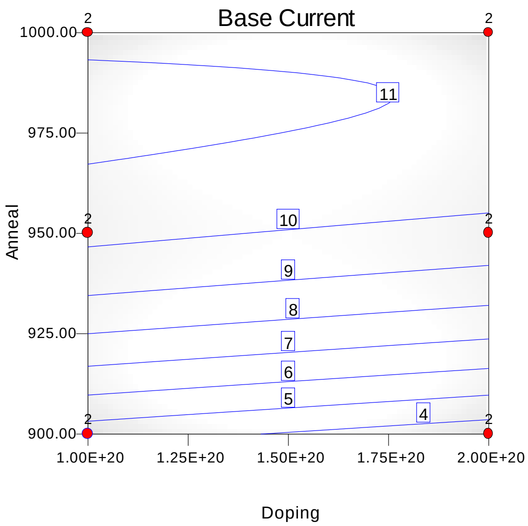
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	112.73	4	28.18	493.73	< 0.0001	significant
A	*0.98*	*1*	*0.98*	*17.18*	*0.0043*	
B	*93.16*	*1*	*93.16*	*1632.09*	*< 0.0001*	
B^2	*18.03*	*1*	*18.03*	*315.81*	*< 0.0001*	
AB	*0.56*	*1*	*0.56*	*9.84*	*0.0164*	
Residual	0.40	7	0.057			
Lack of Fit	*0.014*	*1*	*0.014*	*0.22*	*0.6569*	*not significant*
Pure Error	*0.39*	*6*	*0.064*			
Cor Total	113.13	11				

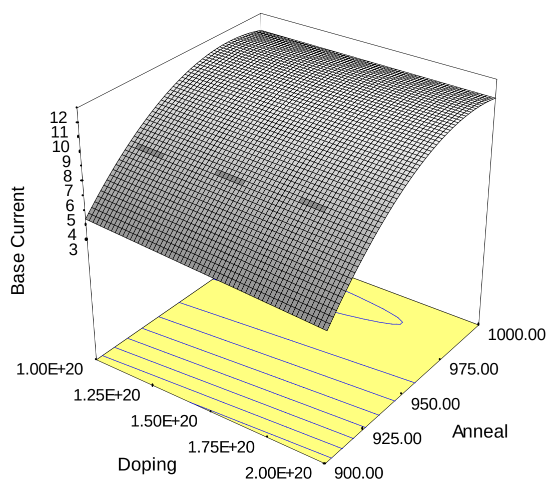
The Model F-value of 493.73 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, B^2, AB are significant model terms.

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	9.94	1	0.12	9.66	10.22	
A-Doping	-0.29	1	0.069	-0.45	-0.12	1.00
B-Anneal	3.41	1	0.084	3.21	3.61	1.00
B^2	-2.60	1	0.15	-2.95	-2.25	1.00
AB	0.27	1	0.084	0.065	0.46	1.00

All of the coefficients in the assumed model are significant. The quadratic effect is easily observable in the response surface plot.





5.26. An experiment was conducted to study the life (in hours) of two different brands of batteries in three different devices (radio, camera, and portable DVD player). A completely randomized two-factor experiment was conducted, and the following data resulted.

Brand of Battery	Device: Radio	Camera	DVD Player
A	8.6	7.9	5.4
	8.2	8.4	5.7
B	9.4	8.5	5.8
	8.8	8.9	5.9

(a) Analyze the data and draw conclusions, using $\alpha = 0.05$.

Both brand of battery (*A*) and type of device (*B*) are significant, the interaction is not.

Design Expert Output

Response: Life
ANOVA for Selected Factorial Model
Analysis of variance table [Terms added sequentially (first to last)]

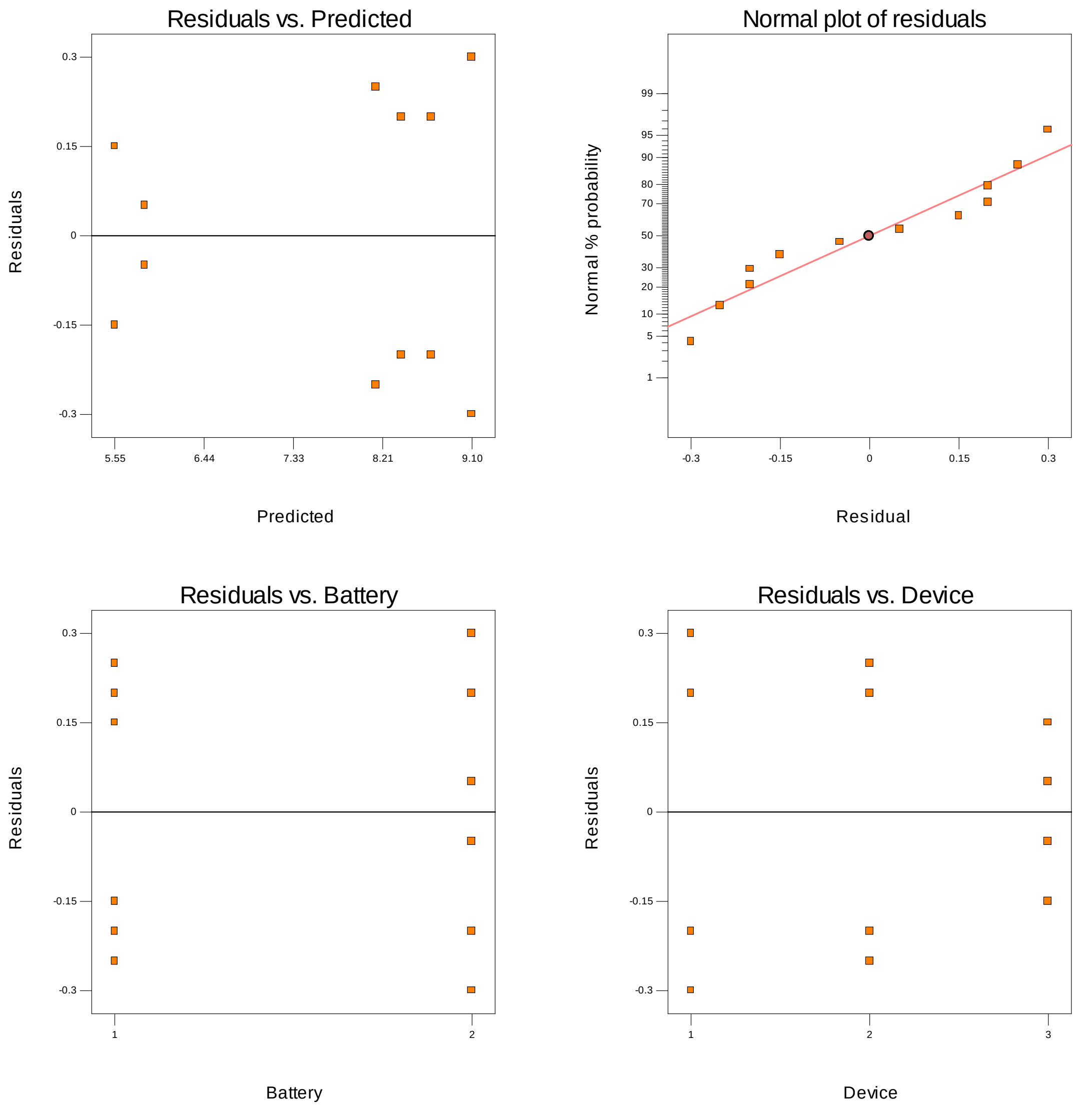
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	23.33	5	4.67	54.36	< 0.0001	significant
A	*0.80*	*1*	*0.80*	*9.33*	*0.0224*	
B	*22.45*	*2*	*11.22*	*130.75*	*< 0.0001*	
AB	*0.082*	*2*	*0.041*	*0.48*	*0.6430*	
Pure Error	*0.52*	*6*	*0.086*			
Cor Total	23.84	11				

The Model F-value of 54.36 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

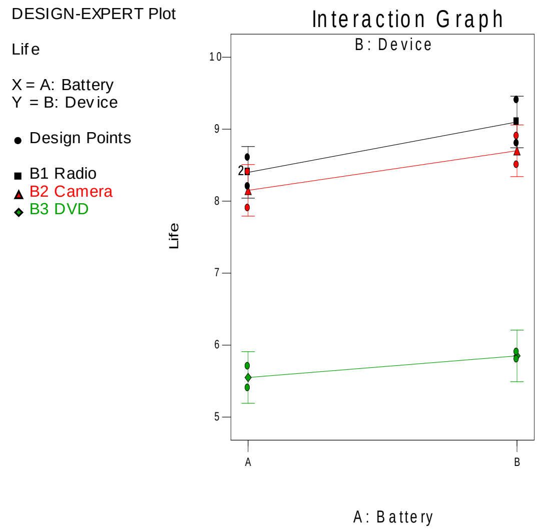
(b) Investigate model adequacy by plotting the residuals.

The residual plots show no serious deviations from the assumptions.



(c) Which brand of batteries would you recommend?

Battery brand B is recommended.



5.27. I have recently purchased new golf clubs, which he believes will significantly improve my game. Below are the scores of three rounds of golf played at three different golf courses with the old and the new clubs.

Clubs	Course		
	Ahwatukee	Karsten	Foothills
Old	90	91	88
	87	93	86
	86	90	90
New	88	90	86
	87	91	85
	85	88	88

(a) Conduct an anlysis of variance. Using $\alpha = 0.05$, what conclusions can you draw?

Although there is a significant difference between the golf courses, there is not a significant difference between the old and new clubs.

Design Expert Output

Response: Score
ANOVA for Selected Factorial Model
Analysis of variance table [Terms added sequentially (first to last)]

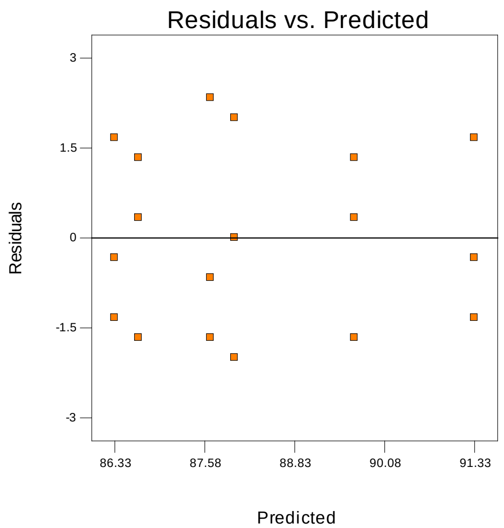
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	54.28	5	10.86	3.69	0.0297	significant
A	*9.39*	*1*	*9.39*	*3.19*	*0.0994*	
B	*44.44*	*2*	*22.22*	*7.55*	*0.0075*	
AB	*0.44*	*2*	*0.22*	*0.075*	*0.9277*	
Pure Error	*35.33*	*12*	*2.94*			
Cor Total	89.61	17				

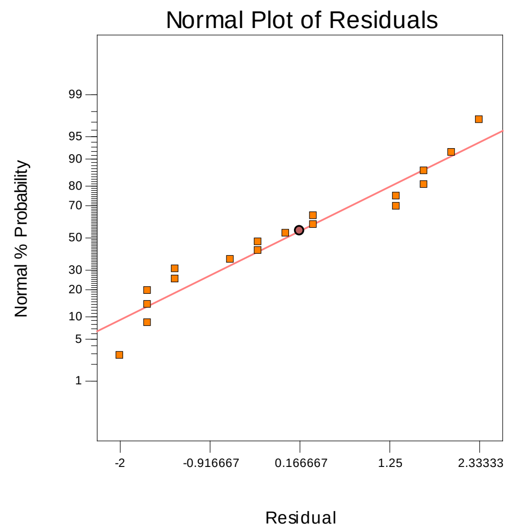
The Model F-value of 3.69 implies the model is significant. There is only a 2.97% chance that a "Model F-Value" this large could occur due to noise.

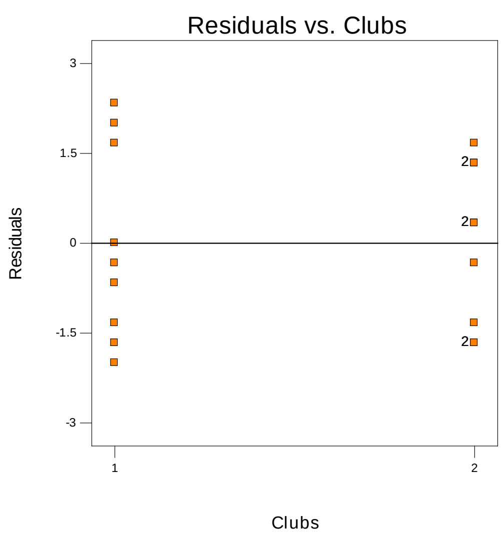
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case B are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

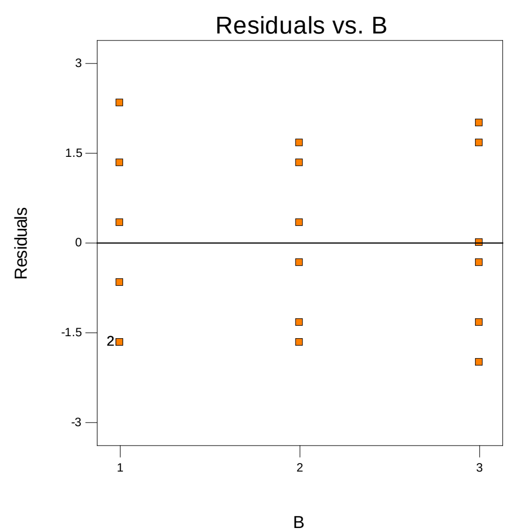
(b) Investigate model adequacy by plotting the residuals.

The residual plots show no deviations from the assumptions.









5.28. A manufacturer of laundry products is investigating the performance of a newly formulated stain remover. The new formulation is compared to the original formulation with respect to its ability to remove a standard tomato-like stain in a test article of cotton cloth using a factorial experiment. The other factors in the experiment are the number of times the test article is washed (1 or 2), and whether or not a detergent booster is used. The response variable is the stain shade after washing (12 is the darkest, 0 is the lightest). The data are shown in the table below.

Formulation	Number of Washings		Number of Washings	
	1	2	1	2
	Booster		Booster	
	Yes	No	Yes	No
New	6	6	3	4
	5	5	2	1
Original	10	11	10	9
	9	11	9	10

(a) Conduct an anlysis of variance. Using $\alpha = 0.05$, what conclusions can you draw?

The formulation, number of washings, and the interaction between these to factors appear to be significant. Continued analysis is required as a result of the residual plots in part (b). Conclusions are presented in part (b).

Design Expert Output

Response: Stain Shade
ANOVA for Selected Factorial Model
Analysis of variance table [Terms added sequentially (first to last)]

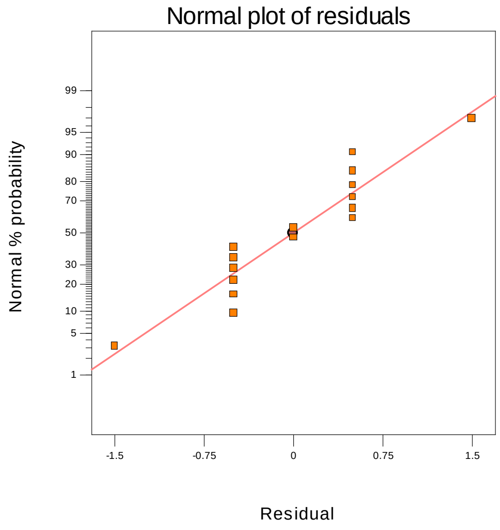
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	159.44	7	22.78	24.30	< 0.0001	significant
A	*138.06*	*1*	*138.06*	*147.27*	*< 0.0001*	
B	*14.06*	*1*	*14.06*	*15.00*	*0.0047*	
C	*0.56*	*1*	*0.56*	*0.60*	*0.4609*	
AB	*5.06*	*1*	*5.06*	*5.40*	*0.0486*	
AC	*0.56*	*1*	*0.56*	*0.60*	*0.4609*	
BC	*0.56*	*1*	*0.56*	*0.60*	*0.4609*	
ABC	*0.56*	*1*	*0.56*	*0.60*	*0.4609*	
Pure Error	*7.50*	*8*	*0.94*			
Cor Total	166.94	15				

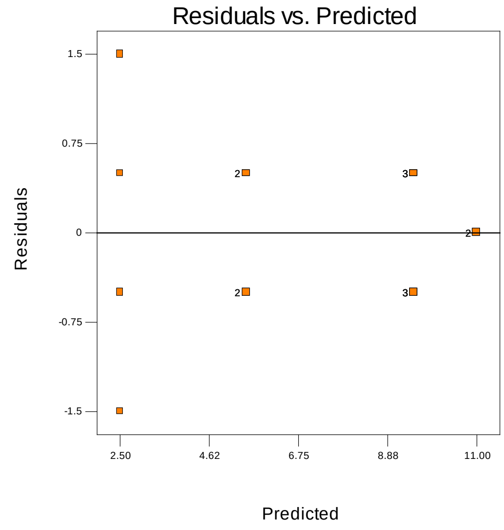
The Model F-value of 24.30 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

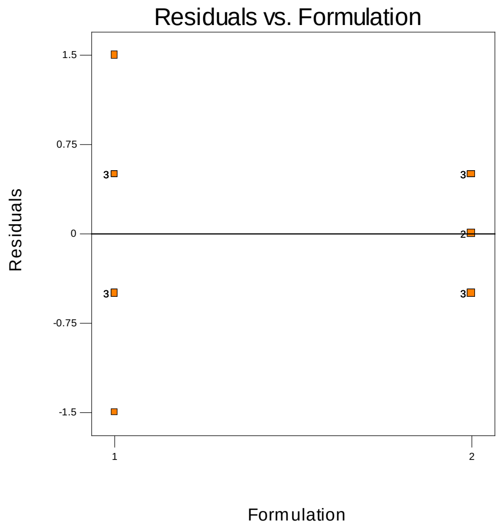
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, AB are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

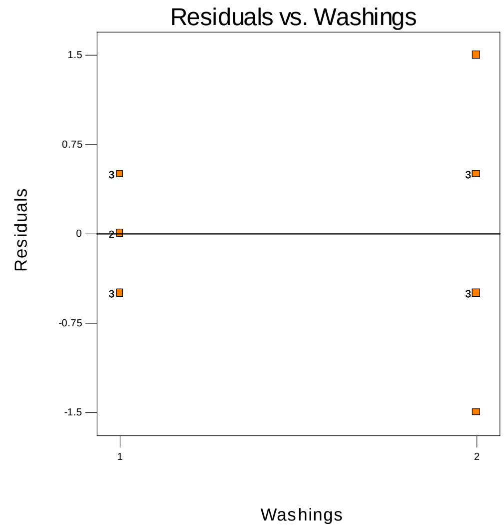
(b) Investigate model adequacy by plotting the residuals.

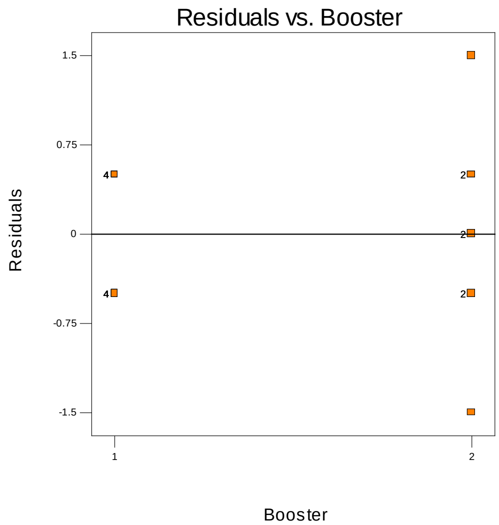
The residual plots shown below identify a violation from our assumptions; nonconstant variance. A power transformation was chosen to correct the violation. λ can be found through trial and error; or the use of a Box-Cox plot that is described in a later chapter. A Box-Cox plot is shown below that identifies a power transformation λ of 1.66.

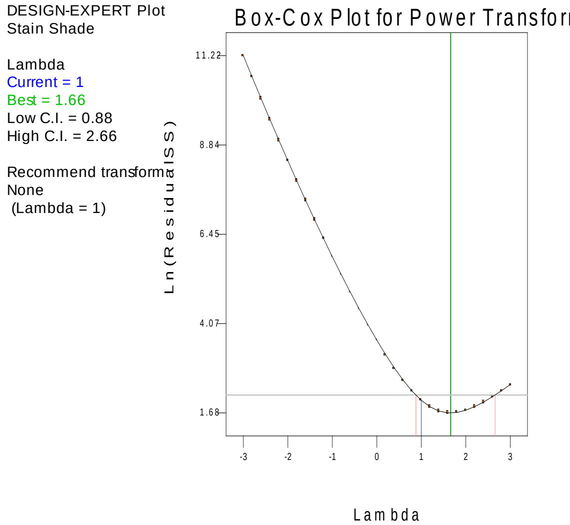

Normal plot of residuals
Normal % probability
99
95
90
80
70
50
30
20
10
5
1
-1.5
-0.75
0
0.75
1.5
Residual


Residuals vs. Predicted
Residuals
1.5
0.75
0
-0.75
-1.5
2.50
4.62
6.75
8.88
11.00
Predicted


Residuals vs. Formulation
Residuals
1.5
0.75
0
-0.75
-1.5
1
2
Formulation


Residuals vs. Washings
Residuals
1.5
0.75
0
-0.75
-1.5
1
2
Washings


Residuals vs. Booster
Residuals
1.5
0.75
0
-0.75
-1.5
1
2
Booster


DESIGN-EXPERT Plot
Stain Shade
Lambda
Current = 1
Best = 1.66
Low C.I. = 0.88
High C.I. = 2.66
Recommend transform
None
(Lambda = 1)
Box-Cox Plot for Power Transforr
Ln(ResidualSS)
11.22
8.84
6.45
4.07
1.68
-3
-2
-1
0
1
2
3
Lambda

The analysis of variance was performed with the transformed data and is shown below. This time, only the formulation and number of washings appear to be significant; the interaction between these two factors is no longer significant after the data transformation. The residual plots show no deviations from the assumptions. The plot of the effects below identfies the new formulation along with two washings produces the best results. The booster is not significant.

Design Expert Output

Response: Stain Shade

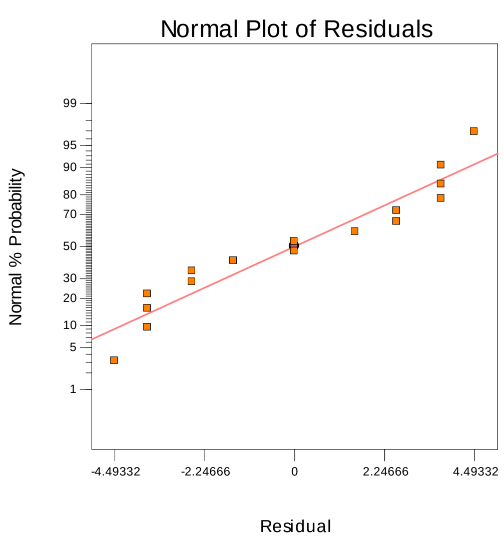
ANOVA for Selected Factorial Model

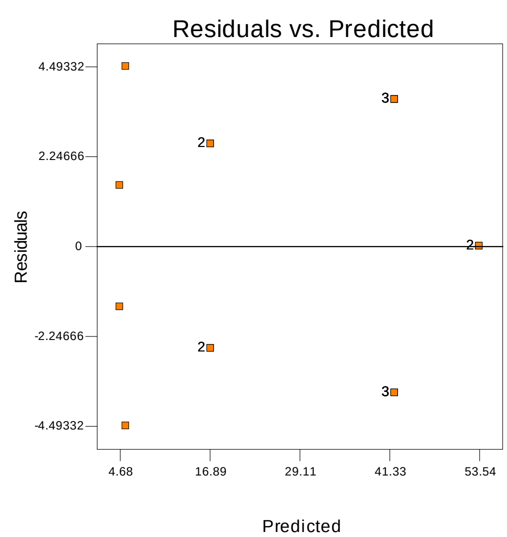
Analysis of variance table [Terms added sequentially (first to last)]

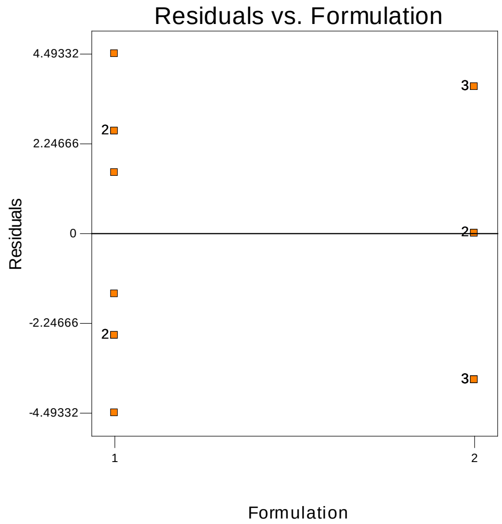
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	5071.22	7	724.46	38.18	< 0.0001	significant
A	*4587.21*	*1*	*4587.21*	*241.74*	*< 0.0001*	
B	*312.80*	*1*	*312.80*	*16.48*	*0.0036*	
C	*37.94*	*1*	*37.94*	*2.00*	*0.1951*	
AB	*38.24*	*1*	*38.24*	*2.01*	*0.1935*	
AC	*28.55*	*1*	*28.55*	*1.50*	*0.2548*	
BC	*28.55*	*1*	*28.55*	*1.50*	*0.2548*	
ABC	*37.94*	*1*	*37.94*	*2.00*	*0.1951*	
Pure Error	*151.81*	*8*	*18.98*			
Cor Total	5223.03	15				

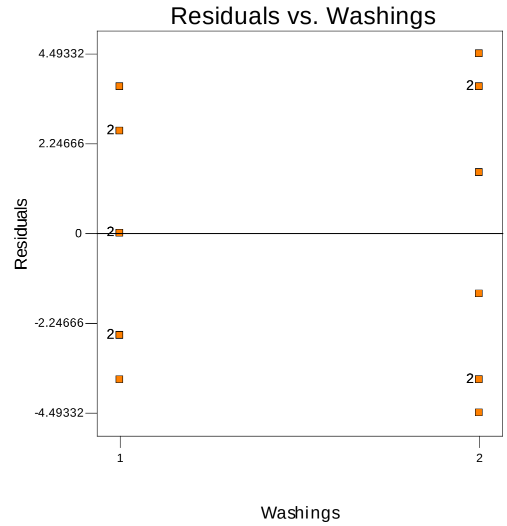
The Model F-value of 38.18 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

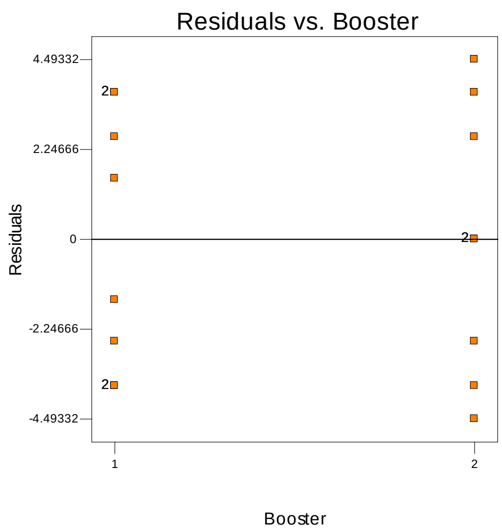
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

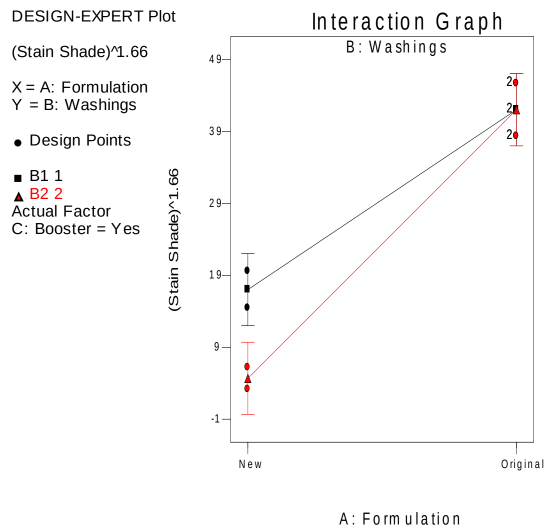












5.29. Bone anchors are used by orthopedic surgeons in repairing torn rotator cuffs (a common shoulder tendon injury among baseball players). The bone anchor is a threaded insert that is screwed into a hole that has been drilled into the shoulder bone near the site of the torn tendon. The torn tendon is then sutured to the anchor. In a successful operation, the tendon is stabilized and reattaches itself to the bone. However, bone anchors can pull out if they are subjected to high loads. An experiment was performed to study the force required to pull out the anchor for three anchor types and two different foam densities (the foam simulates the natural variability found in real bone). Two replicates of the experiment were performed. The experimental design and the pullout force response data are as follows.

	Foam Density			
Anchor Type	Low		High	
A	190	200	241	255
B	185	190	230	237
C	210	205	256	260

(a) Analyze the data from this experiment.

Design Expert Output

Response: **Force**

ANOVA for selected factorial model

Analysis of variance table [Classical sum of squares - Type II]

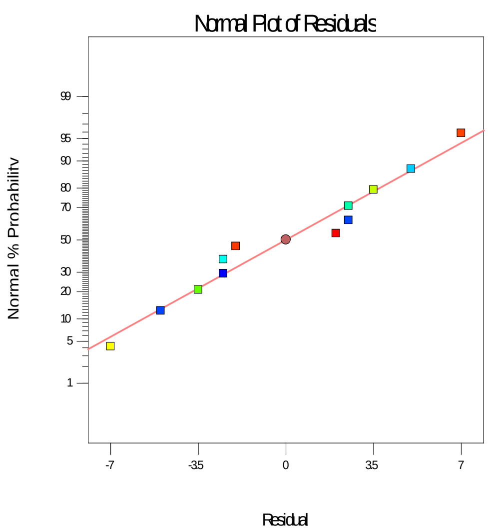
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	8465.42	5	1693.08	49.43	< 0.0001	significant
A-Anchor Type	*990.17*	*2*	*495.08*	*14.45*	*0.0051*	
B-Foam Density	*7450.08*	*1*	*7450.08*	*217.52*	*< 0.0001*	
AB	*25.17*	*2*	*12.58*	*0.37*	*0.7071*	
Pure Error	205.50	6	34.25			
Cor Total	8670.92	11				

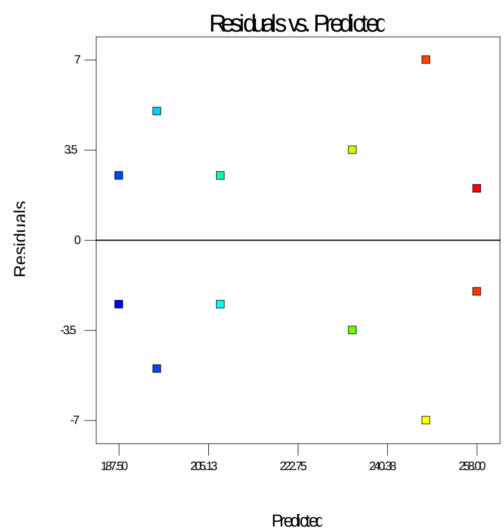
The Model F-value of 49.43 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

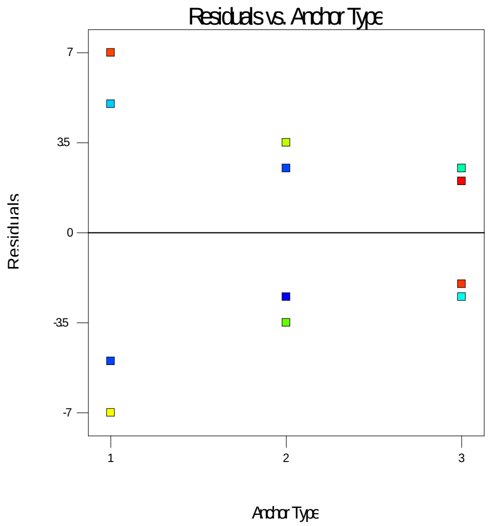
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

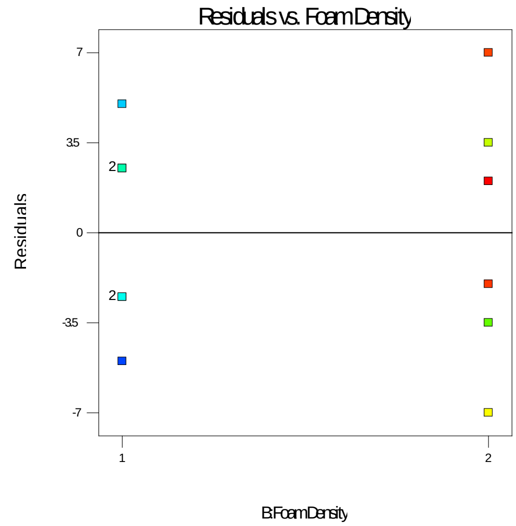
(b) Investigate model adequacy by constructing appropriate residual plots.

The residuals versus Anchor Type appear to show some inequality of variance; however, not enough to be of concern.



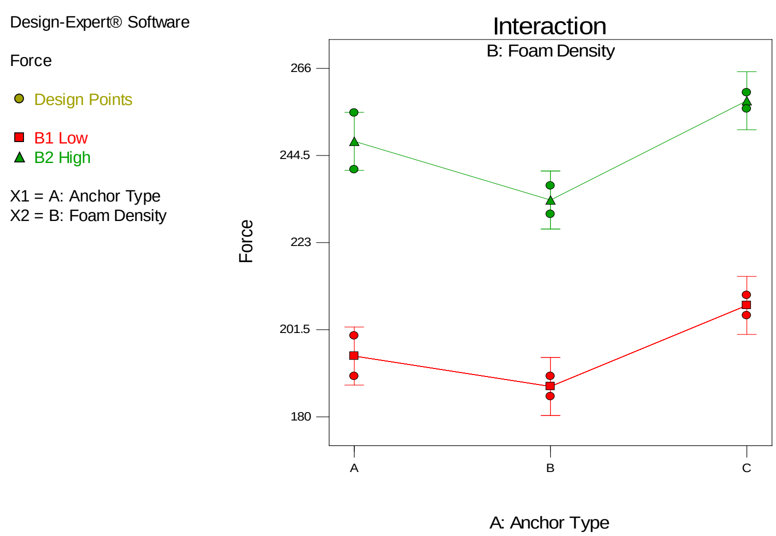






(c) What conclusions can you draw?

Both factors are significant; the interaction is not significant.



5.30. An experiment was performed to investigate the keyboard feel on a computer (crisp or mushy) and the size of the keys (small, medium, or large). The response variable is typing speed. Three replicates of the experiment were performed. The experimental design and the data are as follow.

	Keyboard Feel					
Key Size	Mushy			Crisp		
Small	31	33	35	36	40	41
Medium	36	35	33	40	41	42
Large	37	34	33	38	36	39

(a) Analyze the data from this experiment.

Design Expert Output

Response: Speed

ANOVA for selected factorial model

Analysis of variance table [Classical sum of squares - Type II]

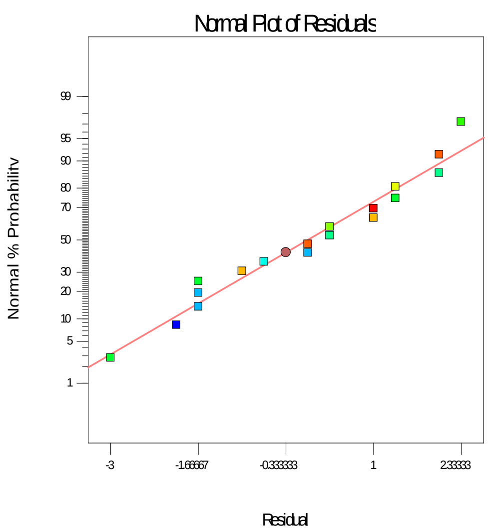
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	140.00	5	28.00	8.00	0.0016	significant
A-Key Size	12.33	2	6.17	1.76	0.2134	
B-Keyboard Feel	117.56	1	117.56	33.59	< 0.0001	
AB	10.11	2	5.06	1.44	0.2741	
Pure Error	42.00	12	3.50			
Cor Total	182.00	17				

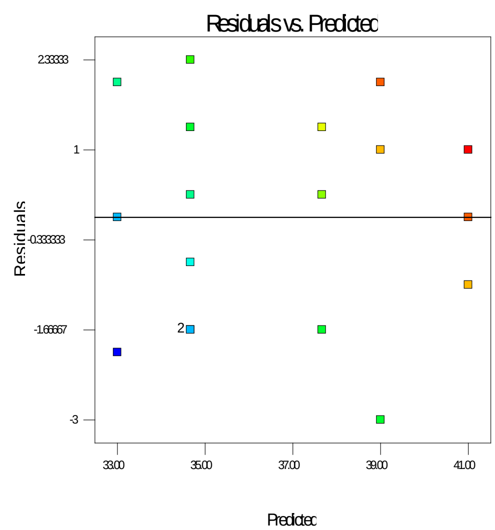
The Model F-value of 8.00 implies the model is significant. There is only a 0.16% chance that a "Model F-Value" this large could occur due to noise.

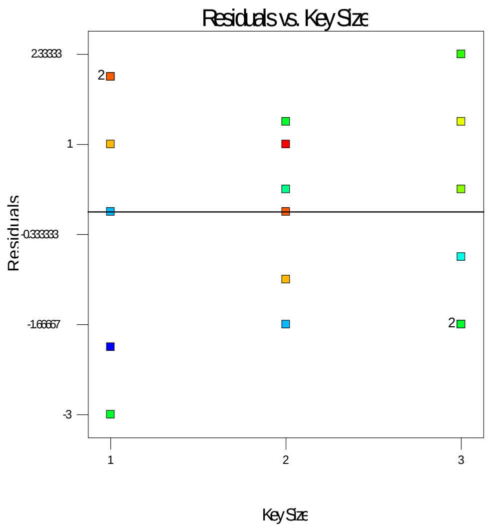
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case B are significant model terms.
Values greater than 0.1000 indicate the model terms are not significant.
If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

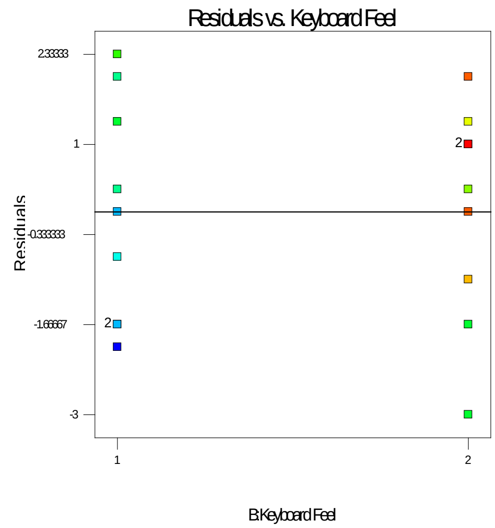
(b) Investigate model adequacy by constructing appropriate residual plots.

The residual plots show no deviations from the assumptions.



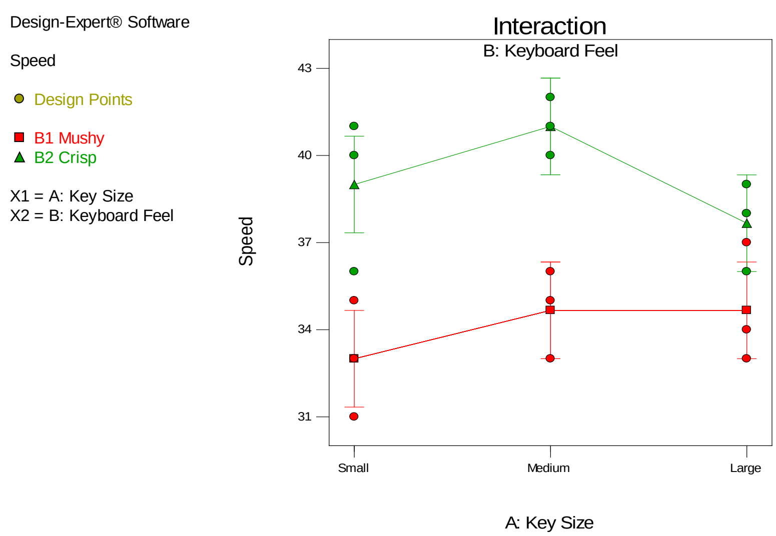






(c) What conclusions can you draw?

Only factor *B*, Keyboard Feel, appears to be significant. From the plot below, the faster typing speed is observed with the crisp keyboard feel.



5.31. An article in *Quality Progress* (May 2011, pp.42-48) describes the use of factorial experiments to improve a silver powder productin process. This product is used in conductive pastes to manufacture a wide variety of products ranging from silicon wafers to elastic membrane switches. Powder density (g/cm^2) and surface area (cm^2/g) are the two critical characteristics of this product. The experiments involved three factors – reaction temperature, ammonium percent, and stirring rate. Each of these factors had two levels and the design was replicated twice. The design is shown below.

Ammonium (%)	Stir Rate (RPM)	Temperature (°C)	Density	Surface Area
2	100	8	14.68	0.40
2	100	8	15.18	0.43
30	100	8	15.12	0.42
30	100	8	17.48	0.41

2	150	8	7.54	0.69
2	150	8	6.66	0.67
30	150	8	12.46	0.52
30	150	8	12.62	0.36
2	100	40	10.95	0.58
2	100	40	17.68	0.43
30	100	40	12.65	0.57
30	100	40	15.96	0.54
2	150	40	8.03	0.68
2	150	40	8.84	0.75
30	150	40	14.96	0.41
30	150	40	14.96	0.41

(a) Analyze the density response. Are any interactions significant? Draw appropriate conclusions about the effects of the significant factors on the response.

The *AB* interaction is significant, along with factors *A* and *B*. See ANOVA table below.

Design Expert Output

Response 1 Density
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	155.19	7	22.17	5.58	0.0136	significant
A-Ammonium	*44.39*	*1*	*44.39*	*11.18*	*0.0102*	
B-Stir Rate	*70.69*	*1*	*70.69*	*17.80*	*0.0029*	
C-Temperature	*0.33*	*1*	*0.33*	*0.083*	*0.7812*	
AB	*28.12*	*1*	*28.12*	*7.08*	*0.0288*	
AC	*0.022*	*1*	*0.022*	*5.480E-003*	*0.9428*	
BC	*10.13*	*1*	*10.13*	*2.55*	*0.1489*	
ABC	*1.52*	*1*	*1.52*	*0.38*	*0.5534*	
Pure Error	31.76	8	3.97			
Cor Total	186.95	15				

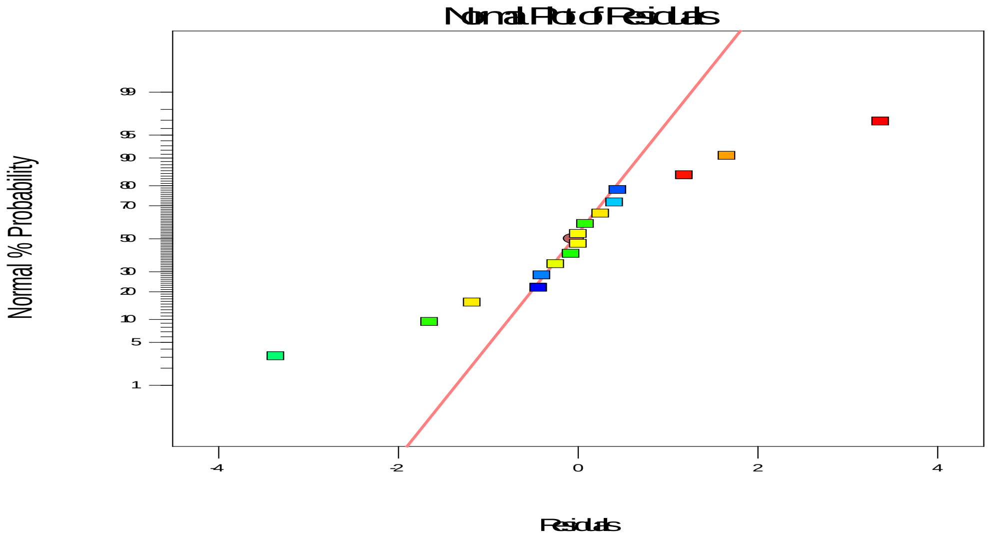
The Model F-value of 5.58 implies the model is significant. There is only a 1.36% chance that a "Model F-Value" this large could occur due to noise.

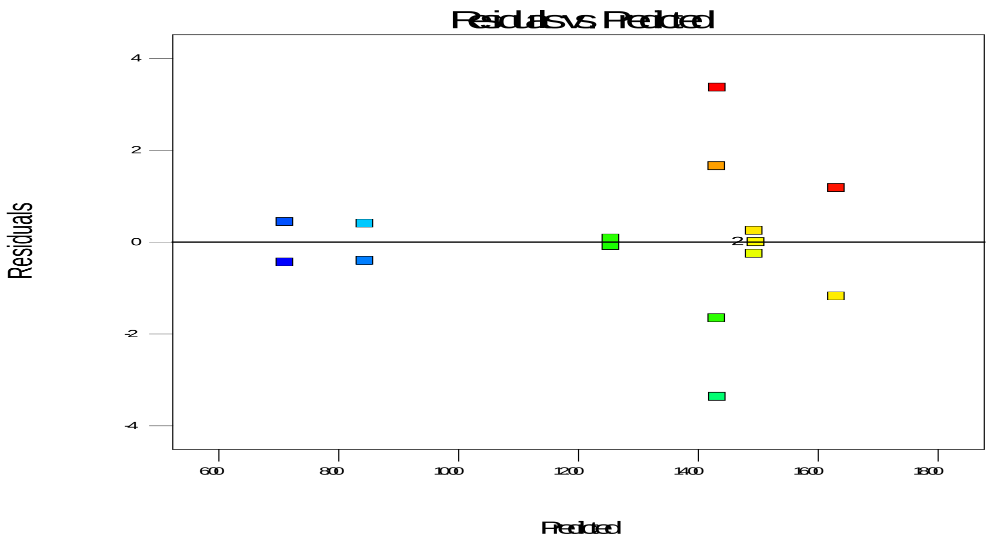
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, AB are significant model terms.

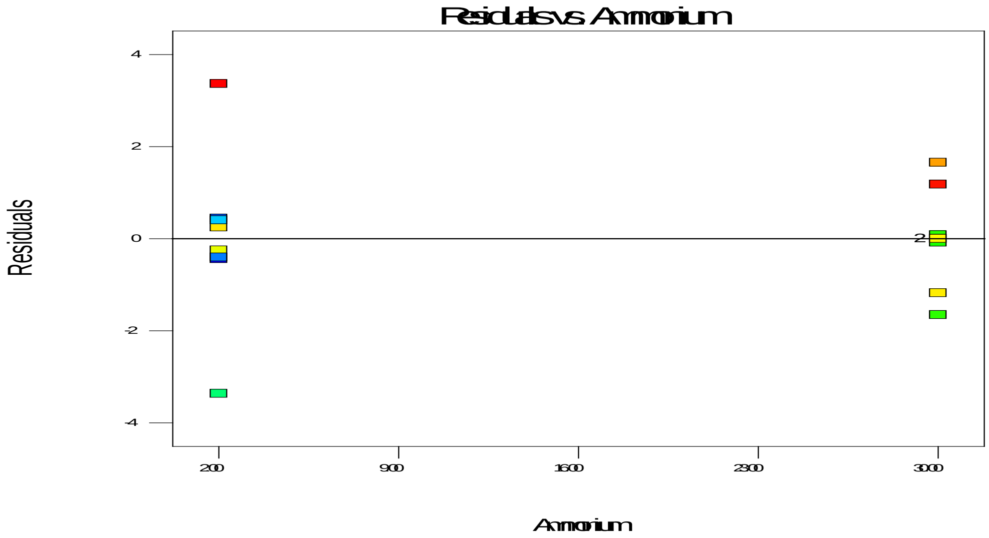
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	12.86	1	0.50	11.71	14.01	
A-Ammonium	1.67	1	0.50	0.52	2.81	1.00
B-Stir Rate	-2.10	1	0.50	-3.25	-0.95	1.00
C-Temperature	0.14	1	0.50	-1.01	1.29	1.00
AB	1.33	1	0.50	0.18	2.47	1.00
AC	-0.037	1	0.50	-1.19	1.11	1.00
BC	0.80	1	0.50	-0.35	1.94	1.00
ABC	0.31	1	0.50	-0.84	1.46	1.00

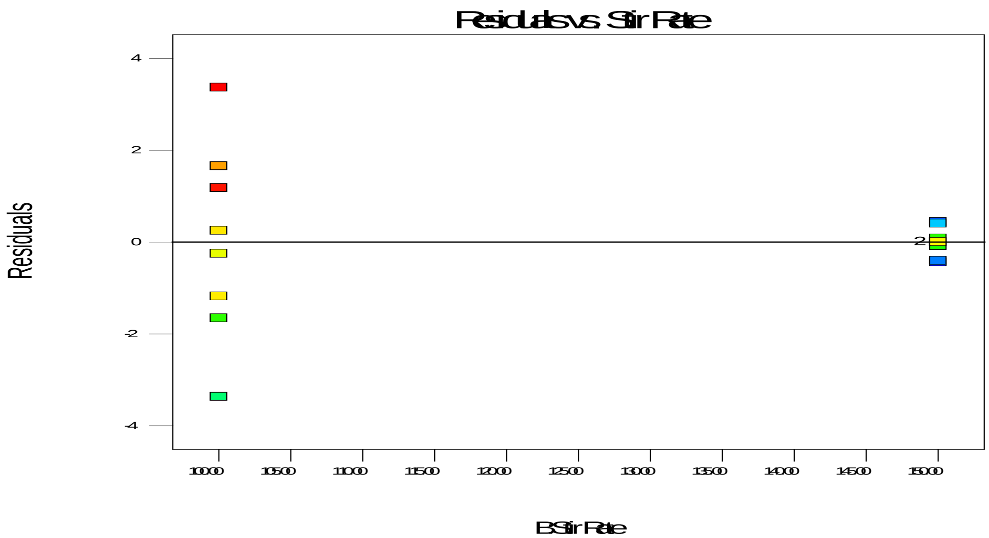
(b) Prepare appropriate residual plots and comment on model adequacy.

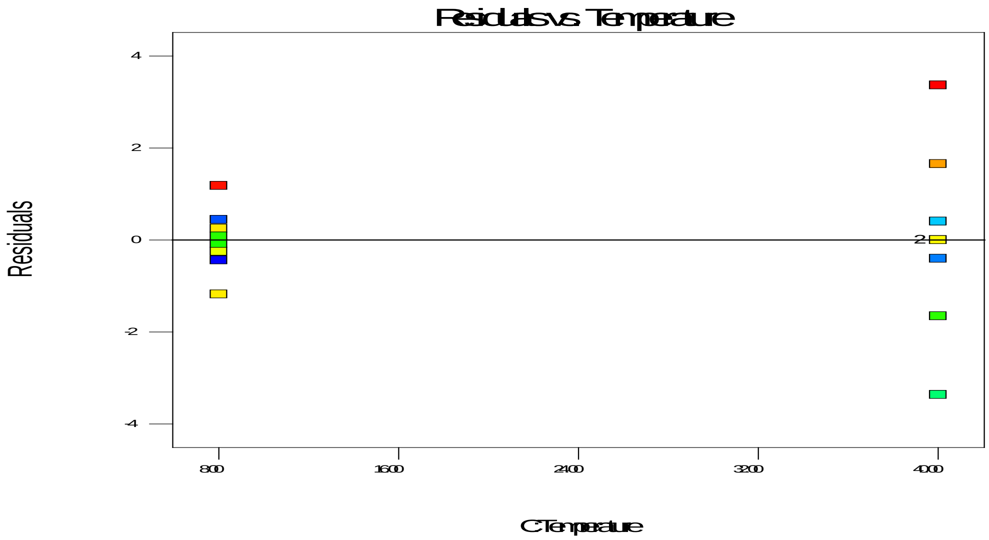
The residual plots below show concerns with the assumptions. A transformation may be appropriate.


Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-4
-2
0
2
4
Residuals


Residuals vs. Predicted
Residuals
4
2
0
-2
-4
2
60.0
80.0
100.0
120.0
140.0
160.0
180.0
Predicted


Residuals vs. Ammonium
Residuals
Ammonium
4
2
0
-2
-4
200
900
1600
2300
3000
2


Residuals vs. Stir Rate
Residuals
B:Stir Rate
4
2
0
-2
-4
1000
1050
1100
1150
1200
1250
1300
1350
1400
1450
1500
2



The inverse transformation was applied to the density data and the analysis and residual plots are shown below.

Design Expert Output

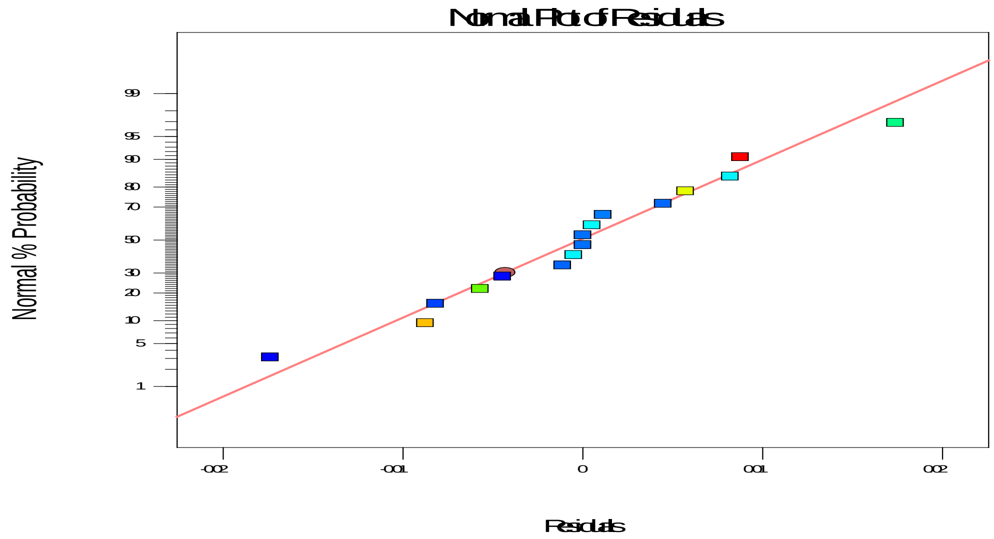
Response 1 Density
Transform: Inverse
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

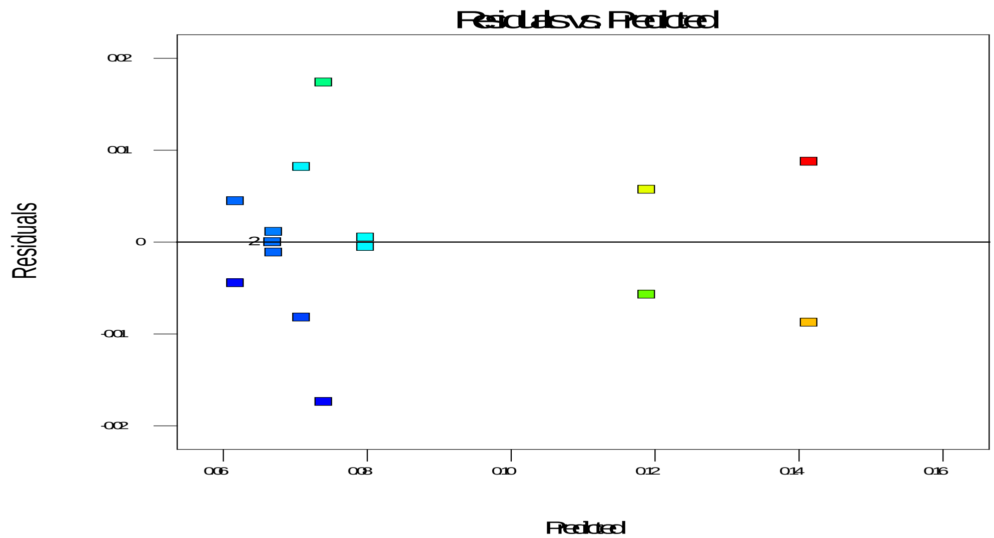
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	0.012	7	1.678E-003	13.42	0.0008	significant
A-Ammonium	*3.723E-003*	*1*	*3.723E-003*	*29.78*	*0.0006*	
B-Stir Rate	*4.445E-003*	*1*	*4.445E-003*	*35.55*	*0.0003*	
C-Temperature	*9.349E-005*	*1*	*9.349E-005*	*0.75*	*0.4124*	
AB	*2.767E-003*	*1*	*2.767E-003*	*22.13*	*0.0015*	
AC	*3.537E-005*	*1*	*3.537E-005*	*0.28*	*0.6093*	
BC	*6.654E-004*	*1*	*6.654E-004*	*5.32*	*0.0499*	
ABC	*1.377E-005*	*1*	*1.377E-005*	*0.11*	*0.7485*	
Pure Error	1.000E-003	8	1.250E-004			
Cor Total	0.013	15				

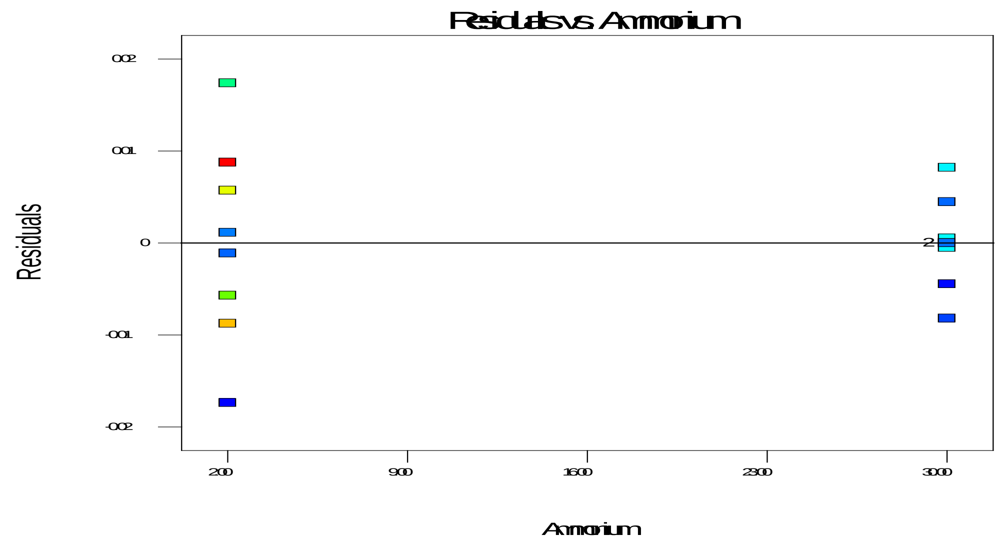
The Model F-value of 13.42 implies the model is significant. There is only a 0.08% chance that a "Model F-Value" this large could occur due to noise.

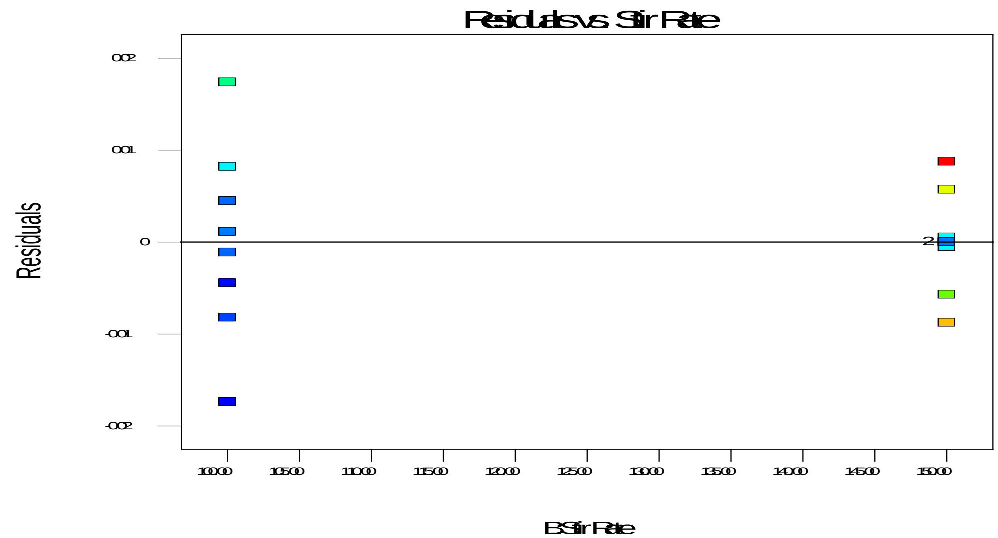
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, AB, BC are significant model terms.

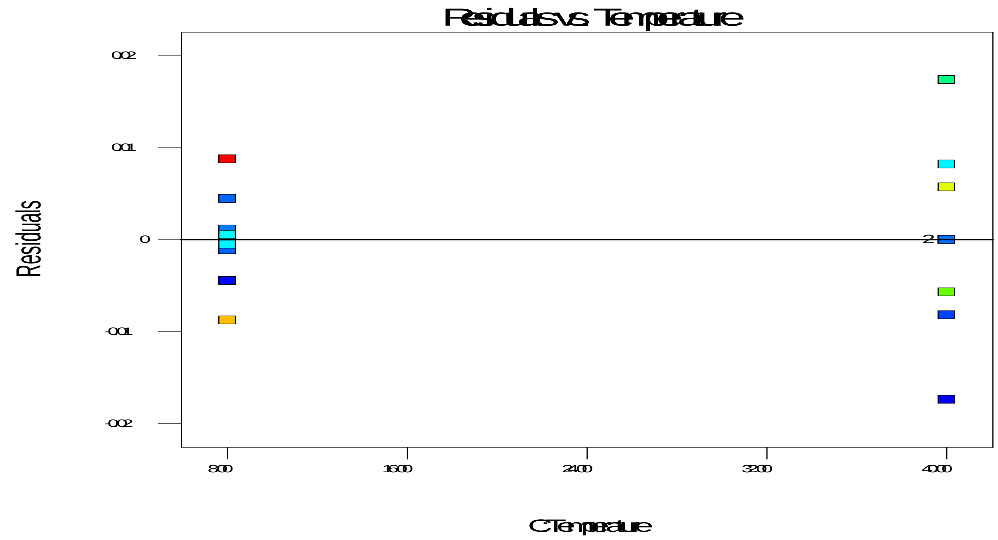
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.085	1	2.795E-003	0.079	0.091	
A-Ammonium	-0.015	1	2.795E-003	-0.022	-8.808E-003	1.00
B-Stir Rate	0.017	1	2.795E-003	0.010	0.023	1.00
C-Temperature	-2.417E-003	1	2.795E-003	-8.863E-003	4.029E-003	1.00
AB	-0.013	1	2.795E-003	-0.020	-6.705E-003	1.00
AC	1.487E-003	1	2.795E-003	-4.959E-003	7.933E-003	1.00
BC	-6.449E-003	1	2.795E-003	-0.013	-2.601E-006	1.00
ABC	9.276E-004	1	2.795E-003	-5.518E-003	7.374E-003	1.00


Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
0
Residuals

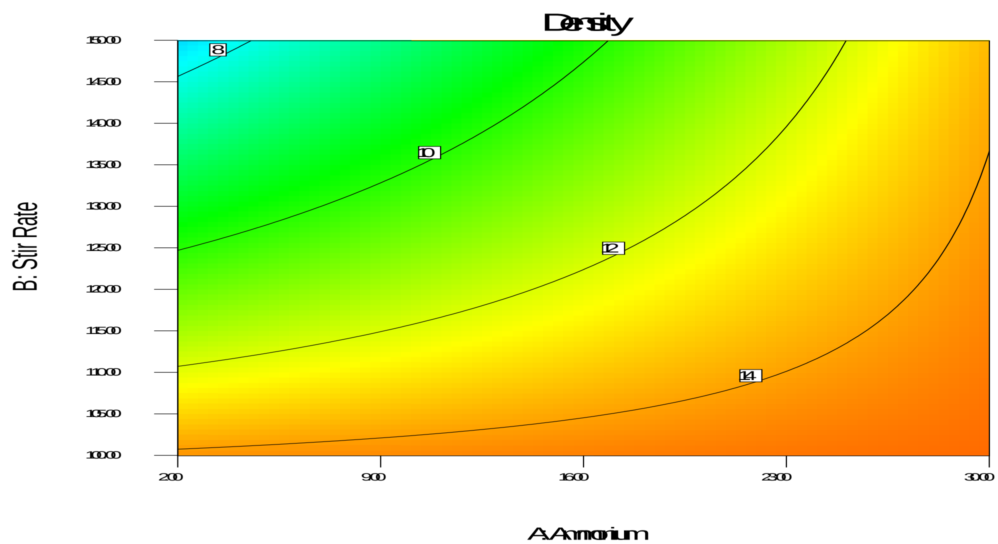

Residuals vs. Predicted
Residuals
0
2
Predicted


Residuals


Residuals



(c) Construct contour plots to aid in practical interpretation of the density response.



(d) Analyze the surface area response. Are any interactions significant? Draw appropriate conclusions about the effects of the significant factors on the response.

The *A* and *B* factors are significant along with the *AB* interaction.

Design Expert Output

Response **2** **Surface Area**

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	0.21	7	0.030	8.73	0.0033	significant
A-Ammonium	*0.061*	*1*	*0.061*	*17.72*	*0.0030*	
B-Stir Rate	*0.032*	*1*	*0.032*	*9.12*	*0.0166*	
C-Temperature	*0.014*	*1*	*0.014*	*3.99*	*0.0807*	

AB	*0.089*	*1*	*0.089*	*25.61*	*0.0010*
AC	*5.625E-005*	*1*	*5.625E-005*	*0.016*	*0.9016*
BC	*0.013*	*1*	*0.013*	*3.66*	*0.0920*
ABC	*3.306E-003*	*1*	*3.306E-003*	*0.96*	*0.3567*
Pure Error	0.028	8	3.456E-003		
Cor Total	0.24	15			

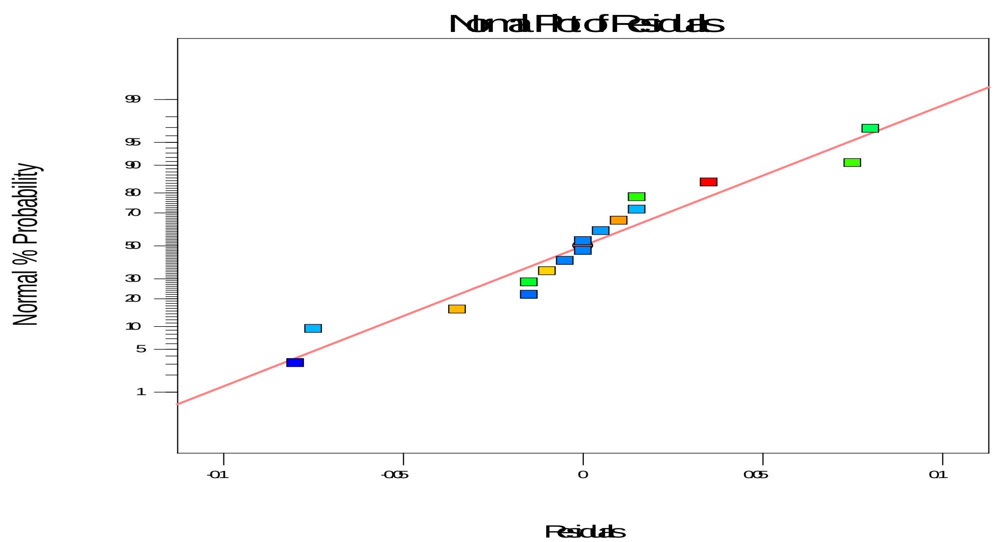
The Model F-value of 8.73 implies the model is significant. There is only
a 0.33% chance that a "Model F-Value" this large could occur due to noise.

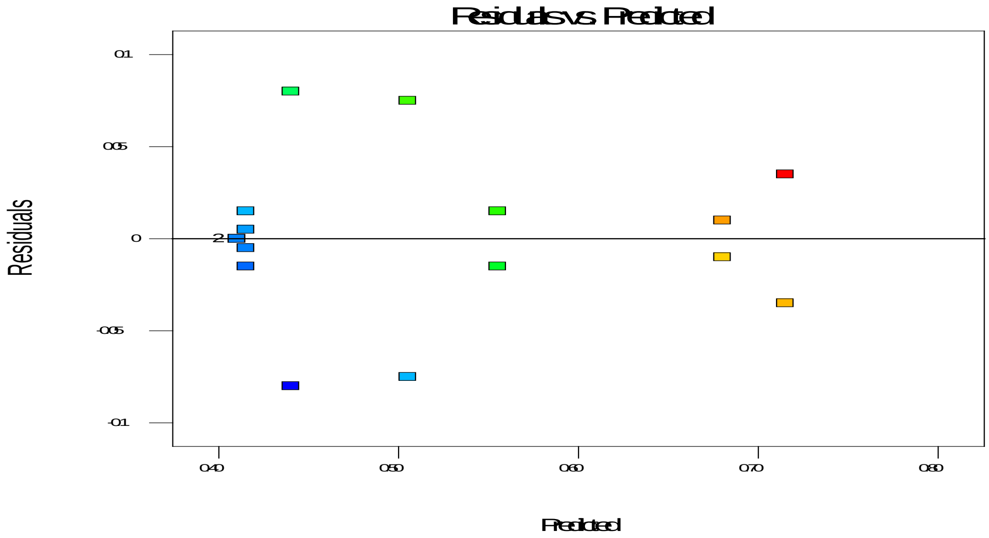
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, AB are significant model terms.

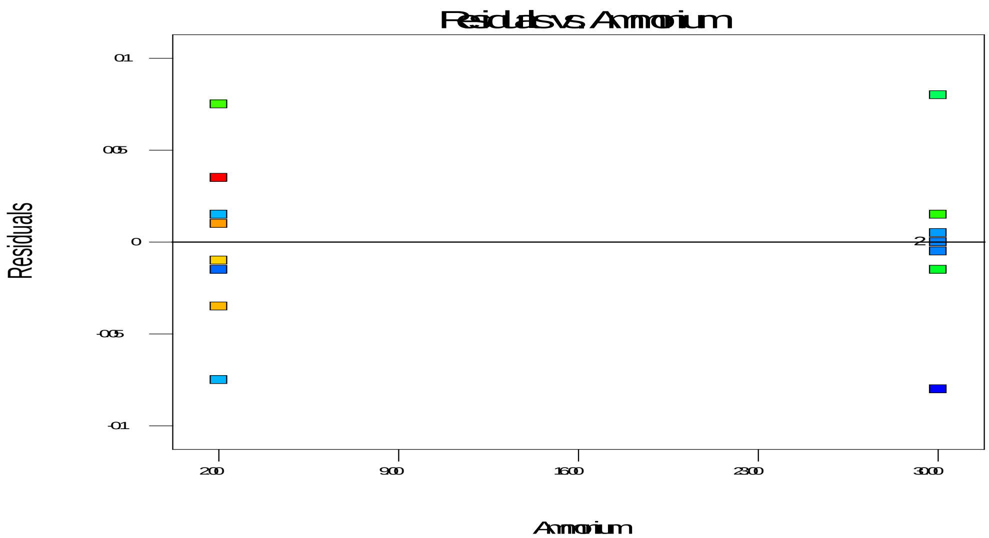
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.52	1	0.015	0.48	0.55	
A-Ammonium	-0.062	1	0.015	-0.096	-0.028	1.00
B-Stir Rate	0.044	1	0.015	0.010	0.078	1.00
C-Temperature	0.029	1	0.015	-4.517E-003	0.063	1.00
AB	-0.074	1	0.015	-0.11	-0.040	1.00
AC	-1.875E-003	1	0.015	-0.036	0.032	1.00
BC	-0.028	1	0.015	-0.062	5.767E-003	1.00
ABC	-0.014	1	0.015	-0.048	0.020	1.00

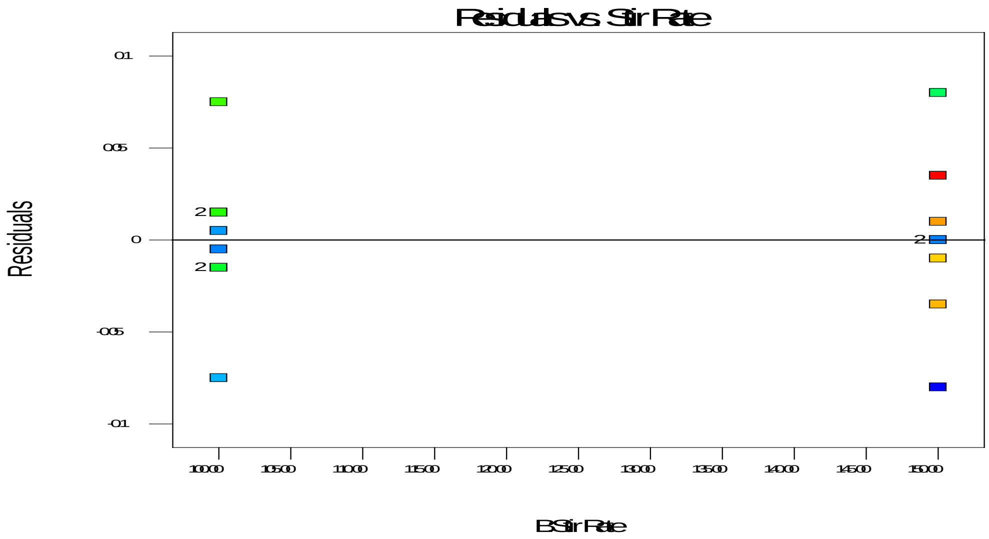
(e) Prepare appropriate residual plots and comment on model adequacy.

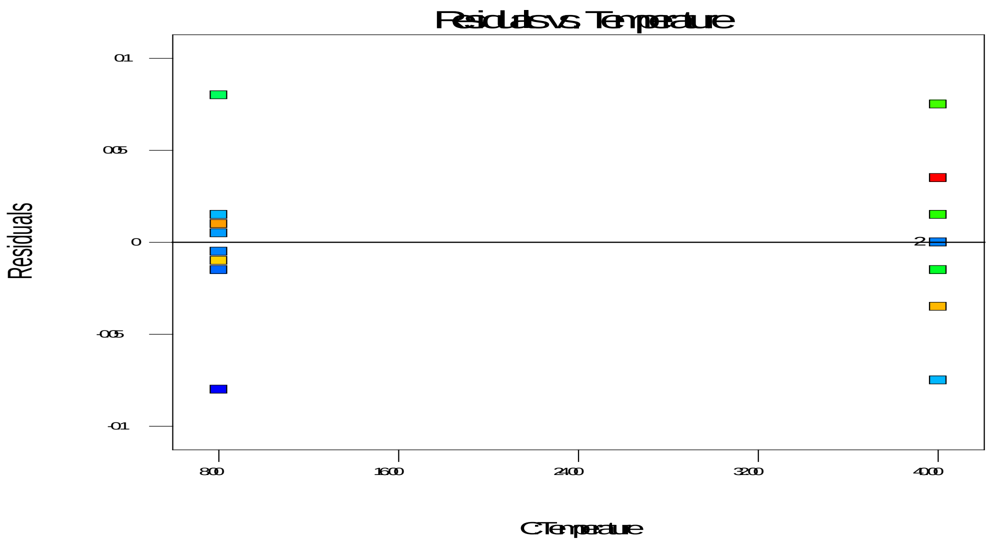
The residual plots below do not identify any concerns with model adequacy.



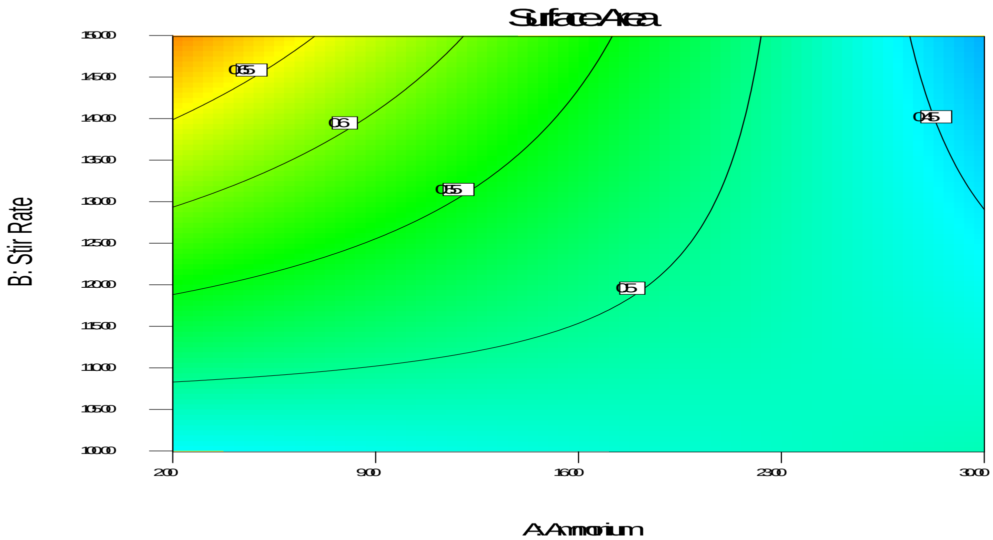

Residuals vs Predicted
Residuals
Predicted


Residuals vs Ammonium
Residuals
Ammonium

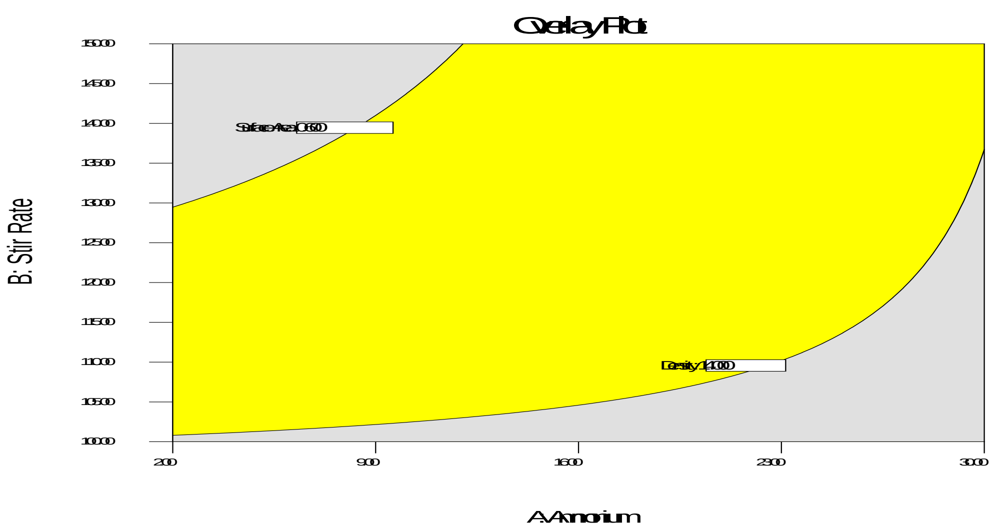




(f) Construct contour plots to aid in practical interpretation of the surface area response.



5.32. Continuation of Problem 5.31. Suppose that the specifications require that surface area must be between 0.3 and 0.6 cm^2/g and that density must be less than 14 g/cm^2. Find a set of operating conditions that will result in a product that meets these requirements.



5.33. An article in *Biotechnology Progress* (2001, Vol. 17, pp. 366-368) described an experiment to investigate nisin extraction in aqueous two-phase solutions. A two-factor factorial experiment was conducted using factors A = concentration of PEG and B = concentration of Na_2SO_4. Data similar to that reported in the paper are shown below.

A	*B*	Extraction (%)
13	11	62.9
13	11	65.4
15	11	76.1
15	11	72.3

13	13	87.5
13	13	84.2
15	13	102.3
15	13	105.6

(a) Analyze the extraction response. Draw appropriate conclusions about the effects of the significant factors on the response.

Factors A and B are significant. The AB interaction is moderately significant.

Response 1 Extraction
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

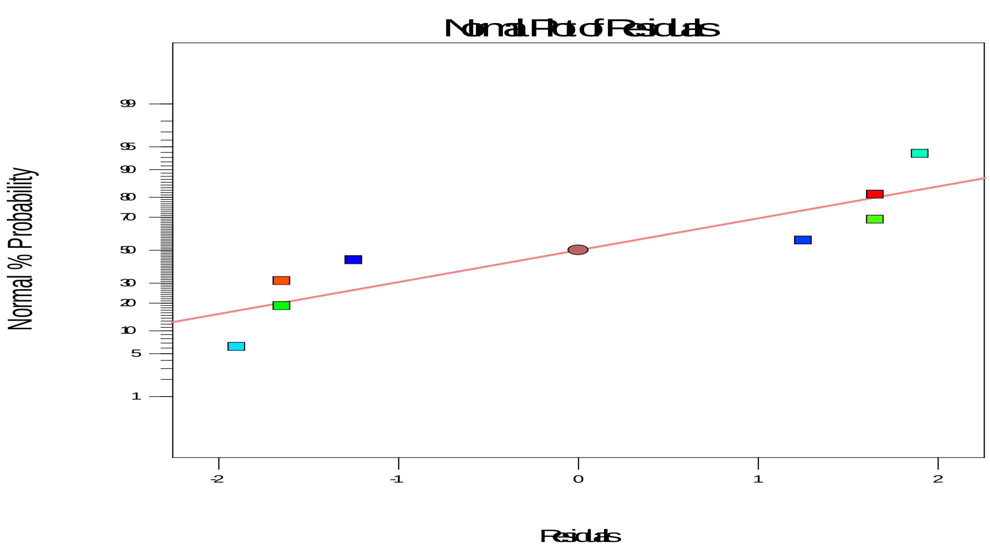
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	1752.16	3	584.05	110.02	0.0003	significant
A-PEG	*396.21*	*1*	*396.21*	*74.63*	*0.0010*	
B-Na2SO4	*1323.55*	*1*	*1323.55*	*249.32*	*< 0.0001*	
AB	*32.40*	*1*	*32.40*	*6.10*	*0.0689*	
Pure Error	21.24	4	5.31			
Cor Total	1773.40	7				

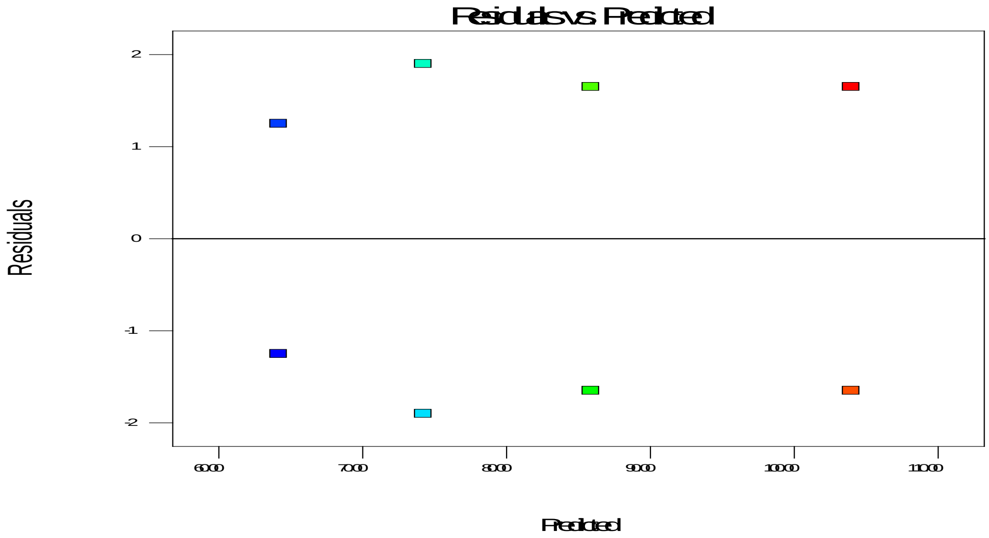
The Model F-value of 110.02 implies the model is significant. There is only a 0.03% chance that a "Model F-Value" this large could occur due to noise.

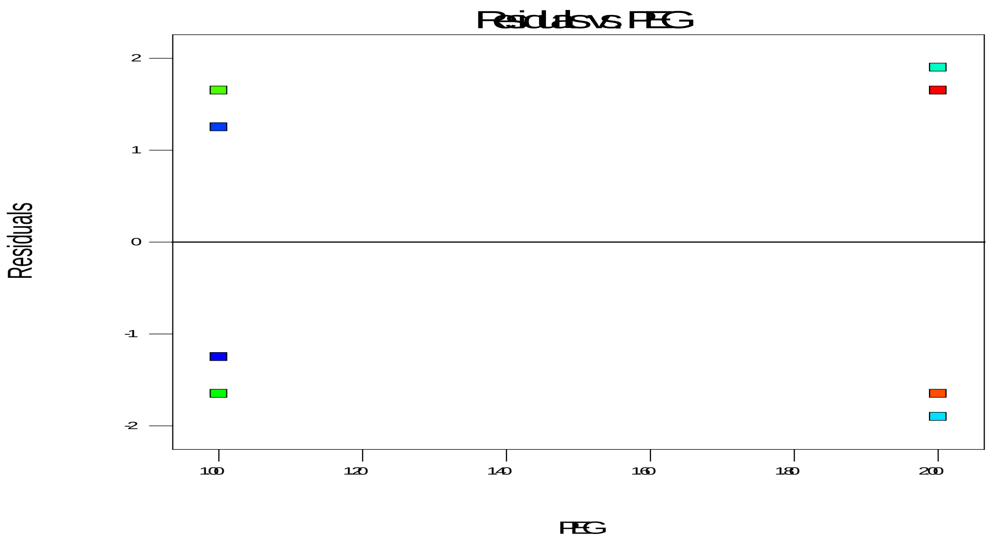
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

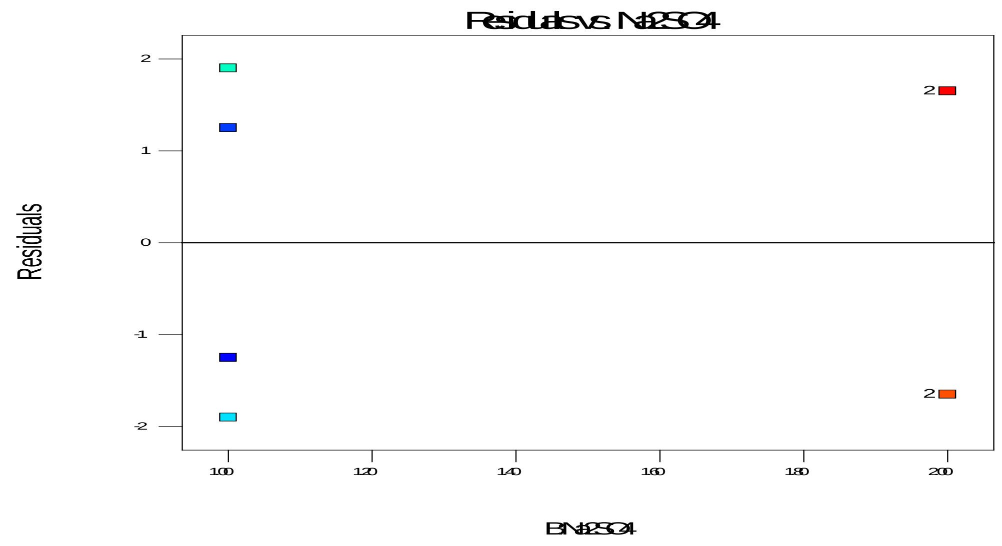
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	82.04	1	0.81	79.78	84.30	
A-PEG	7.04	1	0.81	4.78	9.30	1.00
B-Na2SO4	12.86	1	0.81	10.60	15.12	1.00
AB	2.01	1	0.81	-0.25	4.27	1.00

(b) Prepare appropriate residual plots and comment on model adequacy.

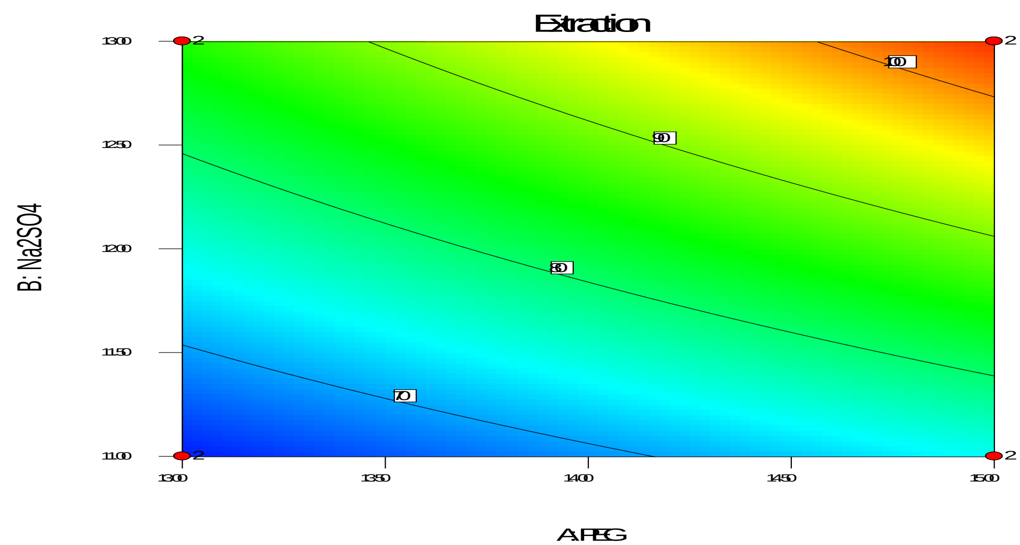



Residuals vs Predicted
Residuals
2
1
0
-1
-2
600
700
800
900
1000
1100
Predicted


Residuals vs PEG
Residuals
2
1
0
-1
-2
100
120
140
160
180
200
PEG



(c) Construct contour plots to aid in practical interpretation of the density response.



5.34. Reconsider the experiment in Problem 5.4. Suppose that this experiment had been conducted in three blocks, with each replicate a block. Assume that the observations in the data table are given in order, that is, the first observation in each cell comes from the first replicate, and so on. Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does it appear that blocking was useful in this experiment?

Feed Rate	Block		Depth of Cut (in)		
(in/min)		0.15	0.18	0.20	0.25
	1	74	79	82	99

0.20	2	64	68	88	104
	3	60	73	92	96
	1	92	98	99	104
0.25	2	86	104	108	110
	3	88	88	95	99
	1	99	104	108	114
0.30	2	98	99	110	111
	3	102	95	99	107

The MS_E was reduced from 28.72 to 23.12. This had very little effect on the results. The variance component estimate for the blocks is:

$$\hat{\sigma}_\beta^2 = \frac{\left[MS_{\text{Blocks}} - MS_E\right]}{ab} = \frac{\left[90.33 - 23.12\right]}{(3)(4)} = 5.60$$

Design Expert Output

Response: Surface Finish
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	180.67	2	90.33			
Model	5842.67	11	531.15	22.97	< 0.0001	significant
A-Feed Rate	*3160.50*	*2*	*1580.25*	*68.35*	*< 0.0001*	
B-Depth of Cut	*2125.11*	*3*	*708.37*	*30.64*	*< 0.0001*	
AB	*557.06*	*6*	*92.84*	*4.02*	*0.0073*	
Residual	508.67	22	23.12			
Cor Total	6532.00	35				

5.35. Reconsider the experiment in Problem 5.6. Suppose that this experiment had been conducted in three blocks, with each replicate a block. Assume that the observations in the data table are given in order, that is, the first observation in each cell comes from the first replicate, and so on. Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does it appear that blocking was useful in this experiment?

Glass Type	Block	Phosphor Type 1	2	3
	1	280	300	290
1	2	290	310	285
	3	285	295	290
	1	230	260	220
2	2	235	240	225
	3	240	235	230

The ANOVA below identifies a very small impact by including the blocks in the analysis. In fact, the MS_E actually increases from 52.78 in Problem 5.6 to 62.50 with the inclusion of the blocks due to the reduction of the residual degrees of freedom from 12 to 10. Because the MS_E is greater than the MS_{Blocks}, the variance component estimate for blocks is zero.

Design Expert Output

Response: Current
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	8.33	2	4.17			
Model	15516.67	5	3103.33	49.65	< 0.0001	significant
A-Phosphor Type	*933.33*	*2*	*466.67*	*7.47*	*0.0104*	
B-Glass Type	*14450.00*	*1*	*14450.00*	*231.20*	*< 0.0001*	
AB	*133.33*	*2*	*66.67*	*1.07*	*0.3803*	
Residual	625.00	10	62.50			
Cor Total	16150.00	17				

5.36. Reconsider the experiment in Problem 5.8. Suppose that this experiment had been conducted in two blocks, with each replicate a block. Assume that the observations in the data table are given in order, that is, the first observation in each cell comes from the first replicate, and so on. Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does it appear that blocking was useful in this experiment?

Operator	Block	Machine			
		1	2	3	4
1	1	109	110	108	110
	2	110	115	109	108
2	1	110	110	111	114
	2	112	111	109	112
3	1	116	112	114	120
	2	114	115	119	117

The ANOVA below identifies a very small impact by including the blocks in the analysis. In fact, the MS_E actually increases from 3.79 in Problem 5.8 to 3.95 with the inclusion of the blocks due to the reduction of the residual degrees of freedom from 12 to 11. Because the MS_E is greater than the MS_{Blocks}, the variance component estimate for blocks is zero.

Design Expert Output

Response:Stength
ANOVA for Selected Factorial Model
Analysis of variance table [Clssical sum of squares – Type II]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	2.04	1	2.04			
Model	217.46	11	19.77	5.00	0.0064	significant
A-Operator	*160.33*	*2*	*80.17*	*20.29*	*0.0002*	
B-Machine	*12.46*	*3*	*4.15*	*1.05*	*0.4087*	
AB	*44.67*	*6*	*7.44*	*1.88*	*0.1716*	
Residual	43.46	11	3.95			
Cor Total	262.96	23				

5.37. Reconsider the three-factor factorial experiment in Problem 5.18. Suppose that this experiment had been conducted in two blocks, with each replicate a block. Assume that the observations in the data table given in order, that is, the first observation in each cell comes from the first replicate, and so on.

Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does is appear that blocking was useful in this experiment?

Percentage of Hardwood Concentration	Block	Cooking Time 3.0 Hours			Cooking Time 4.0 Hours		
		Pressure			Pressure		
		400	500	650	400	500	650
2	1	196.6	197.7	199.8	198.4	199.6	200.6
	2	196.0	196.0	199.4	198.6	200.4	200.9
4	1	198.5	196.0	198.4	197.5	198.7	199.6
	2	197.2	196.9	197.6	198.1	198.0	199.0
8	1	197.5	195.6	197.4	197.6	197.0	198.5
	2	196.6	196.2	198.1	198.4	197.8	199.8

The ANOVA below identifies a very small impact by including the blocks in the analysis; the SS_{Blocks} estimate is zero. In fact, the MS_E actually increases from 0.37 in Problem 5.18 to 0.39 with the inclusion of the blocks due to the reduction of the residual degrees of freedom from 18 to 17. Because the MS_E is greater than the MS_{Blocks}, the variance component estimate for blocks is zero.

Response: Strength
ANOVA for Selected Factorial Model
Analysis of variance table [Clssical sum of squares – Type II]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	0.000	1	0.000			
Model	59.73	17	3.51	9.08	< 0.0001	significant
A-Concentration	*7.76*	*2*	*3.88*	*10.03*	*0.0013*	
B-Time	*20.25*	*1*	*20.25*	*52.32*	*< 0.0001*	
C-Pressure	*19.37*	*2*	*9.69*	*25.03*	*< 0.0001*	
AB	*2.08*	*2*	*1.04*	*2.69*	*0.0967*	
AC	*6.09*	*4*	*1.52*	*3.93*	*0.0194*	
BC	*2.19*	*2*	*1.10*	*2.84*	*0.0866*	
ABC	*1.97*	*4*	*0.49*	*1.27*	*0.3186*	
Residual	6.58	17	0.39			
Cor Total	66.31	35				

5.38. Reconsider the three-factor factorial experiment in Problem 5.19. Suppose that this experiment had been conducted in three blocks, with each replicate a block. Assume that the observations in the data table given in order, that is, the first observation in each cell comes from the first replicate, and so on. Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does is appear that blocking was useful in this experiment?

		Temperature					
		300°C			350°C		
		Operator			Operator		
Cycle Time	Block	1	2	3	1	2	3
	1	23	27	31	24	38	34
40	2	24	28	32	23	36	36
	3	25	26	29	28	35	39
	1	36	34	33	37	34	34
50	2	35	38	34	39	38	36
	3	36	39	35	35	36	31
	1	28	35	26	26	36	28
60	2	24	35	27	29	37	26

3	27	34	25		25	34	24

The ANOVA below identifies a very small impact by including the blocks in the analysis. The MS_E improved from 3.28 in Problem 5.19 to 3.27 with the inclusion of the blocks. The variance component estimate for the blocks is:

$$\hat{\sigma}_\beta^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{abc} = \frac{[3.39 - 3.27]}{(3)(3)(2)} = 0.0067$$

Design Expert Output

Response: Score
ANOVA for Selected Factorial Model
Analysis of variance table [Clssical sum of squares – Type II]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	6.78	2	3.39			
Model	1239.33	17	72.90	22.29	< 0.0001	significant
A-Cycle Time	*436.00*	*2*	*218.00*	*66.64*	*< 0.0001*	
B-Operator	*261.33*	*2*	*130.67*	*39.94*	*< 0.0001*	
C-Temperature	*50.07*	*1*	*50.07*	*15.31*	*0.0004*	
AB	*355.67*	*4*	*88.92*	*27.18*	*< 0.0001*	
AC	*78.81*	*2*	*39.41*	*12.05*	*0.0001*	
BC	*11.26*	*2*	*5.63*	*1.72*	*0.1941*	
ABC	*46.19*	*4*	*11.55*	*3.53*	*0.0163*	
Residual	111.22	34	3.27			
Cor Total	1357.33	53				

5.39. Reconsider the bone anchor experiment in Problem 5.29. Suppose that this experiment had been conducted in two blocks, with each replicate a block. Assume that the observations in the data table are given in order, that is, the first observation in each cell comes from the first replicate, and so on. Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does it appear that blocking was useful in this experiment?

	Foam Density			
Anchor Type	Low		High	
A	190	200	241	255
B	185	190	230	237
C	210	205	256	260

The MS_E was reduced from 28.72 to 23.12. This had very little effect on the results. The MS_E was reduced from 34.25 in Problem 5.29 to 20.68 with the inclusion of the blocks. This had a small effect on the results. The variance component estimate for the blocks is:

$$\hat{\sigma}_\beta^2 = \frac{[MS_{\text{Blocks}} - MS_E]}{ab} = \frac{[102.08 - 20.68]}{(3)(2)} = 13.57$$

Design Expert Output

Response: Force
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Block	102.08	1	102.08			
Model	8465.42	5	1693.08	81.86	< 0.0001	significant
A-Anchor Type	*990.17*	*2*	*495.08*	*23.94*	*0.0028*	

B-Foam Density	*7450.08*	*1*	*7450.08*	*360.20*	*< 0.0001*
AB	*25.17*	*2*	*12.58*	*0.61*	*0.5801*
Residual	103.42	5	20.68		
Cor Total	8670.92	11			

5.40. Reconsider the keyboard experiment in Problem 5.30. Suppose that this experiment had been conducted in three blocks, with each replicate a block. Assume that the observations in the data table are given in order, that is, the first observation in each cell comes from the first replicate, and so on. Reanalyze the data as a factorial experiment in blocks and estimate the variance component for blocks. Does it appear that blocking was useful in this experiment?

	Keyboard Feel					
Key Size	Mushy			Crisp		
Small	31	33	35	36	40	41
Medium	36	35	33	40	41	42
Large	37	34	33	38	36	39

The ANOVA below identifies a very small impact by including the blocks in the analysis. In fact, the MS_E actually increases from 3.50 in Problem 5.30 to 3.97 with the inclusion of the blocks due to the reduction of the residual degrees of freedom from 12 to 10. Because the MS_E is greater than the MS_{Blocks}, the variance component estimate for blocks is zero.

Design Expert Output

Response: Speed
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Block	2.33	2	1.17			
Model	140.00	5	28.00	7.06	0.0045	significant
A-Key Size	*12.33*	*2*	*6.17*	*1.55*	*0.2583*	
B-Keyboard Feel	*117.56*	*1*	*117.56*	*29.64*	*0.0003*	
AB	*10.11*	*2*	*5.06*	*1.27*	*0.3213*	
Residual	39.67	10	3.97			
Cor Total	182.00	17				

5.41. The C. F. Eye Care company manufactures lenses for transplantation into the eye following cataract surgery. An engineering group has conducted an experiment involving two factors to determine their effect on the lens polishing process. The results of this experiment are summarized in the following ANOVA display:

Source	DF	SS	MS	F	*P*-value
Factor A	---	---	0.0833	0.05	0.952
Factor B	---	96.333	96.3333	57.80	<0.001
Interaction	2	122.167	6.0833	3.65	---
Error	6	10.000	---		
Total	11	118.667			

(a) The sum of squares for factor A is <u>0.1666</u>.
(b) The number of degrees of freedom for factor A in the experiment is <u>2</u>.
(c) The number of degrees of freedom for factor B is <u>1</u>.
(d) The mean square for error is <u>1.666</u>.
(e) An upper bound for the P-value for the interaction test statistic is <u>0.1</u>.
(f) The engineers used <u>3</u> levels of factor A in this experiment.
(g) The engineers used <u>2</u> levels of factor B in this experiment.

(h) There are 2 replicates of this experiment.
(i) Would you conclude that the effect of factor B depends on the level of factor A (Yes or No)? No, the P-value is 0.092.
(j) An estimate of the standard deviation of the response variable is 1.29.

Source	DF	SS	MS	F	*P*-value
Factor A	2	0.1666	0.0833	0.05	0.952
Factor B	1	96.333	96.3333	57.80	<0.001
Interaction	2	122.167	6.0833	3.65	0.092
Error	6	10.000	1.6666		
Total	11	118.667			

5.42. Reconsider the lens polishing experiment in Problem 5.41. Suppose that this experiment has been conducted as a randomized complete box design. The sum of squares for blocks is 4.00. Reconstruct the ANOVA given this new information. What impact does the blocking have on the conclusions from the original experiment?

Source	DF	SS	MS	F	*P*-value
Factor A	2	0.1666	0.0833	0.05	0.952
Factor B	1	96.333	96.3333	57.80	<0.001
Interaction	2	122.167	6.0833	3.65	0.092
Block	1	4.000	4.000	0.666	0.452
Error	5	6.000	1.2000		
Total	11	118.667			

Blocking has no impact.

5.43. In Problem 4.53 you met physics PhD Laura Van Ertia who had conducted a single-factor experiment in her pursuit of the unified theory. She is at it again, and this time she has moved on to a two-factor factorial conducted as a completely randomized design. From her experiment, Laura has constructed the following incomplete ANOVA display:

Source	SS	DF	MS	F
Factor A	350.00	2		
Factor B	300.00		150	
Interaction	200.00		50	
Error	150.00	18		
Total	1000.00			

(a) How many levels of factor B did she use in the experiment? 3
(b) How many degrees of freedom are associated with the interaction? 4
(c) The error mean square is ____. 8.333
(d) The mean square for factor A is ____. 175
(e) How many replicates of the experiment were conducted? 3
(f) What are your conclusions about interaction and the two main effects? All are significant
(g) An estimate of the standard deviation of the response variable is ____. 2.886
(h) If this experiment had been run in blocks ther would have been ____ degrees of freedom for blocks. 2

Source	SS	DF	MS	F	P
Factor A	350.00	2	175	21	0.00001968
Factor B	300.00	2	150	18	0.00005081
Interaction	200.00	4	50	6	0.002996

Error	150.00	18	8.333		
Total	1000.00	26			

5.44. Continuation of Problem 5.43. Suppose that Laura did actually conduct the experiment in Problem 5.43 as a randomized complete block design. Assume that the block sum of squares is 60.00. Reconstruct the ANOVA display under the new set of assumptions.

Source	SS	DF	MS	F	P
Factor A	350.00	2	175	31.1	3.0712E-06
Factor B	300.00	2	150	26.7	7.9815E-06
Interaction	200.00	4	50	8.9	0.0005573
Block	60.00	2	30	5.3	0.01714
Error	90.00	16	5.625		
Total	1000.00	26			

In addition to both main effects and the interaction, the block is significant.

5.45. Consider the following ANOVA for a two-factor experiment:

Source	DF	SS	MS	F	*P*-value
Factor A	2	8.0000	4.00000	2.00	0.216
Factor B	1	8.3333	8.33333	4.17	0.087
Interaction	2	10.6667	5.33333	2.67	0.148
Error	6	12.0000	2.0000		
Total	11	39.0000			

In addition to the ANOVA, you are given the following data totals. Row totals (factor A) = 18, 10, 14; column totals (factor B) = 16, 26; cell totals = 10,8,2,8,4,10; and replicate totals = 19, 23. The grand total is 42. The original experiment was a completely randomized block design. Now suppose that the experiment had been run in two complete blocks. Answer the following questions about the ANOVA for the blocked design.

(a) The block sum of squares is _____.

$$SS_{\text{block}} = \frac{1}{ab}\left(block_1^2 + block_2^2\right) - \frac{42^2}{3x2x2} = \frac{1}{2x3}\left(19^2\ + 23^2\ \right) - 147 = 1.3333$$

(b) There are ____ degrees of freedom for blocks. 2
(c) The error sum of squares is now ____. 10.6667
(d) The interaction effect is now significant at 1 percent (Yes or No). No

	B1	B2	Row Total
A1	10	8	18
A2	2	8	10
A3	4	10	14
Column Total	16	26	

Source	DF	SS	MS	F	*P*-value
Factor A	2	8.0000	4.00000	1.87	0.2475
Factor B	1	8.3333	8.33333	3.91	0.1049
Interaction	2	10.6667	5.33333	2.50	0.1768
Block	1	1.3333	1.333	0.57	0.4843
Error	5	10.6667	2.1333		

Total	11	39.0000			

5.46. Consider the following incomplete ANOVA table:

Source	SS	DF	MS	F
Factor A	50.00	1	50.00	
Factor B	80.00	2	40.00	
Interaction	30.00	2	15.00	
Error		12		
Total	172.00	17		

In addition to the ANOVA table, you know that the experiment has been replicated three times, and that the totals of the three replicates are 10, 12, and 14, repectively. The original experiment was run as a completely randomized design. Answer the following questions:

(a) The pure error estimate of the standard deviation of the sample observations is 1 (Yes or No)?
Yes

Source	SS	DF	MS	F
Factor A	50.00	1	50.00	
Factor B	80.00	2	40.00	
Interaction	30.00	2	15.00	
Error	12.00	12	1.00	
Total	172.00	17		

(b) Suppose that the experiment has been run in blocks, so that it is a randomized complete block design. The number of degrees of freedom for blocks would be ____. 2

(c) The block sum of squares is ____?

$$SS_{\text{block}} = \frac{1}{ab}\left(block_1^2 + block_2^2 + block_3^2\right) - \frac{36^2}{2x3x3} = \frac{1}{2x3}\left(10^2 + 12^2 + 14^2\right) - 72 = 1.3333$$

(d) The error sum of squares in the randomized complete block design is now ____? 10.66

(e) For the randomized complete block design, what is the extimate of the standard deviation of the sample observations? 0.94

Source	SS	DF	MS	F
Factor A	50.00	1	50.00	
Factor B	80.00	2	40.00	
Interaction	30.00	2	15.00	
Block	1.333	2	0.66	
Error	10.66	12	0.89	
Total	172.00	17		

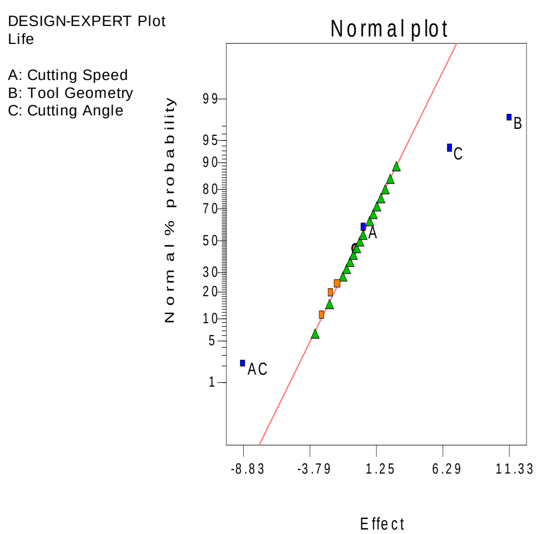
Chapter 6
The 2^k Factorial Design
Solutions

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:

			Treatment	Replicate		
A	*B*	*C*	Combination	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	significant
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	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			
Lack of Fit	*93.00*	*3*	*31.00*	*1.03*	*0.4067*	*not significant*
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.

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

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:

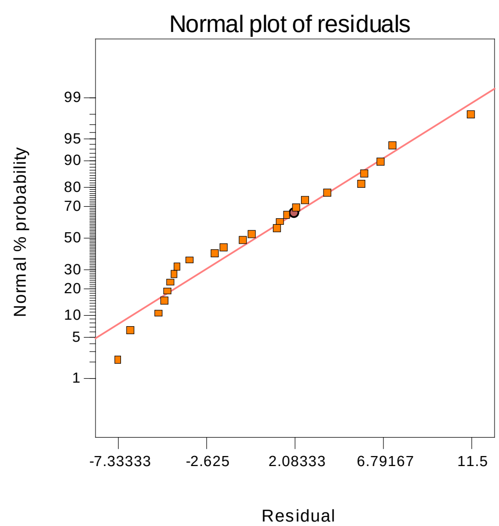
Life	=
+40.83	
+0.17	* A
+5.67	* B
+3.42	* C
-4.42	* A * C

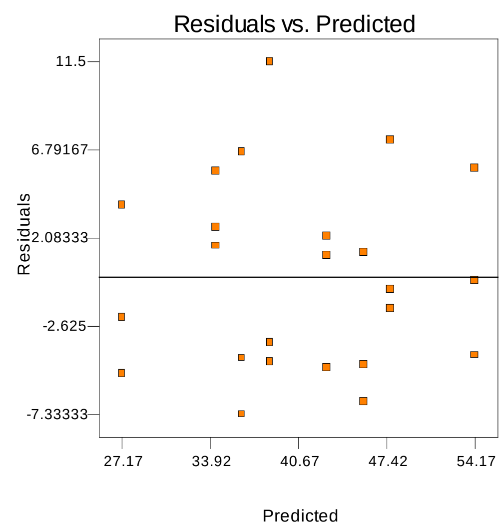
Final Equation in Terms of Actual Factors:

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 "hierarchial" 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?

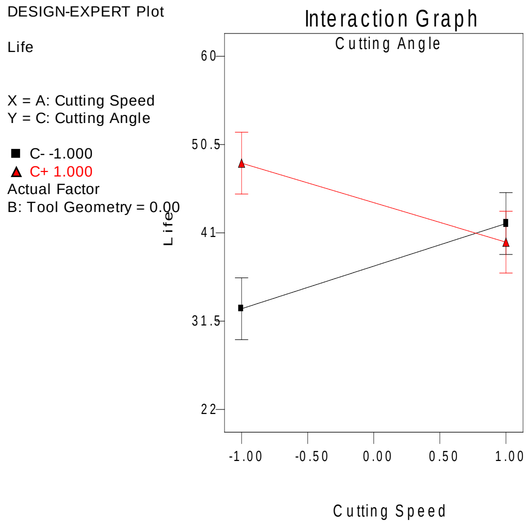


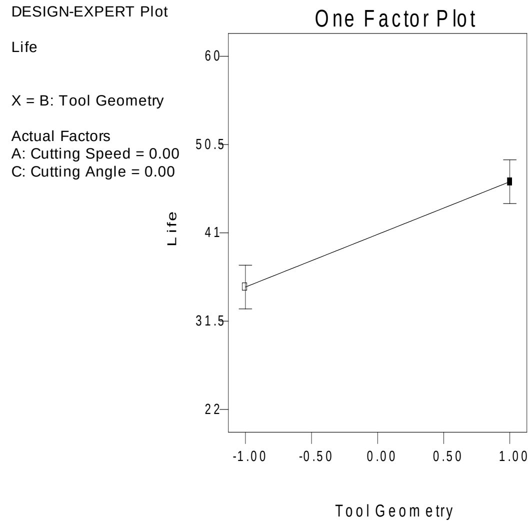


There is nothing unusual about the residual plots.

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

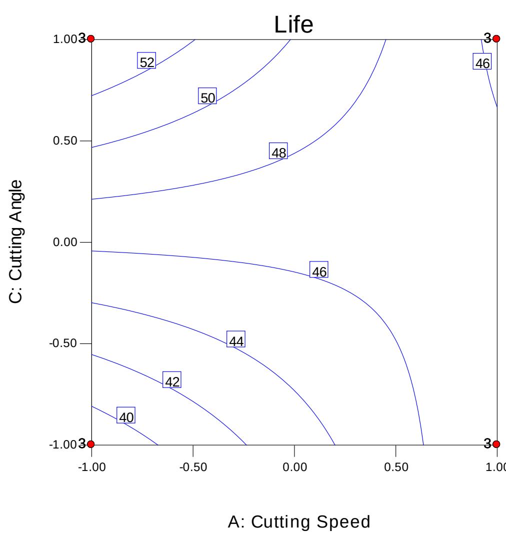
Since *B* has a positive effect, set *B* at the high level to increase life. The *AC* interaction plot reveals that life would be maximized with *C* at the high level and *A* at the low level.

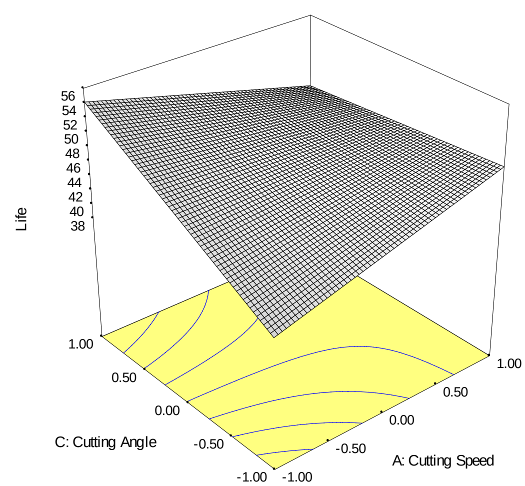




6.2. Reconsider part (c) of Problem 6.1. Use the regression model to generate response surface and contour plots of the tool life response. Interpret these plots. Do they provide insight regarding the desirable operating conditions for this process?

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.





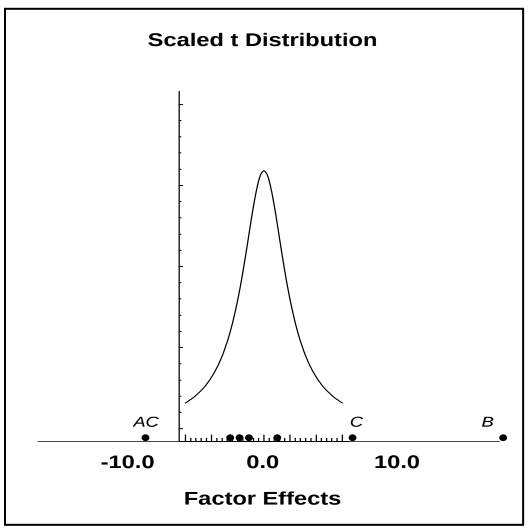
6.3. Find the standard error of the factor effects and approximate 95 percent confidence limits for the factor effects in Problem 6.1. Do the results of this analysis agree with the conclusions from the analysis of variance?

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

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.

		Treatment	Replicate			
A	*B*	Combination	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

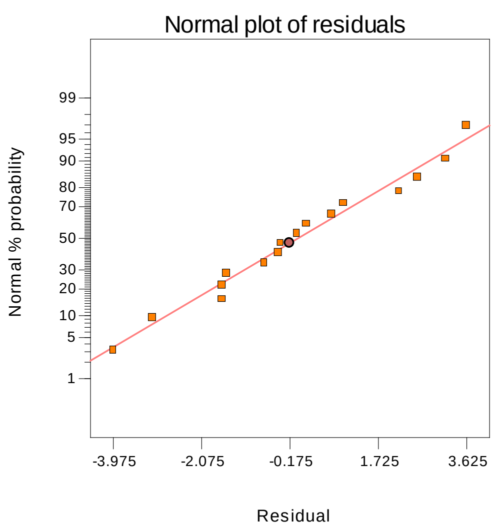
ANOVA for Selected Factorial Model

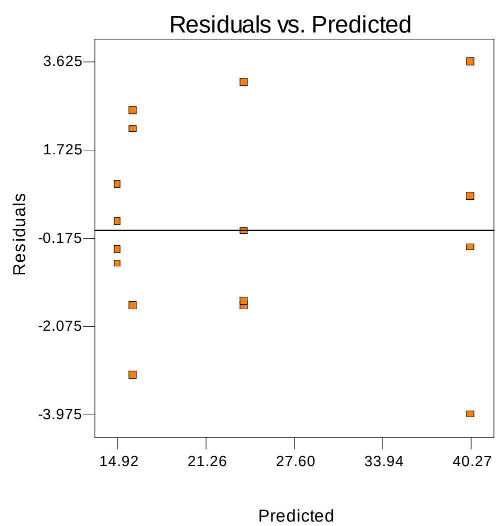
Analysis of variance table [Partial sum of squares]

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.

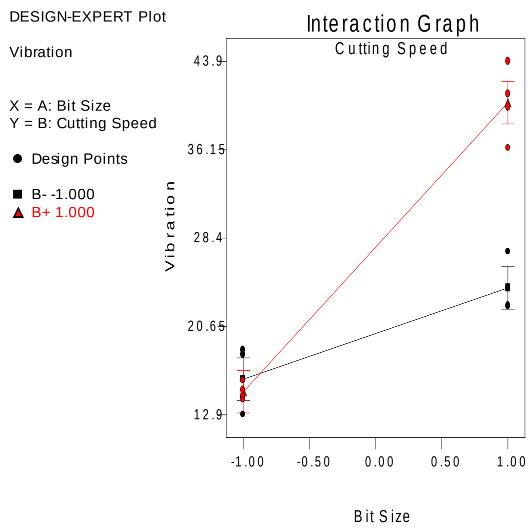




There is nothing unusual about the residual plots.

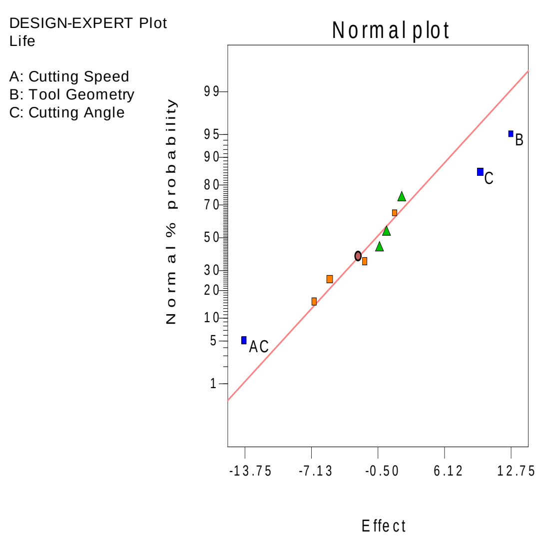
(c) Draw the *AB* interaction plot. Interpret this plot. What levels of bit size and speed would you recommend for routine operation?

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

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)?

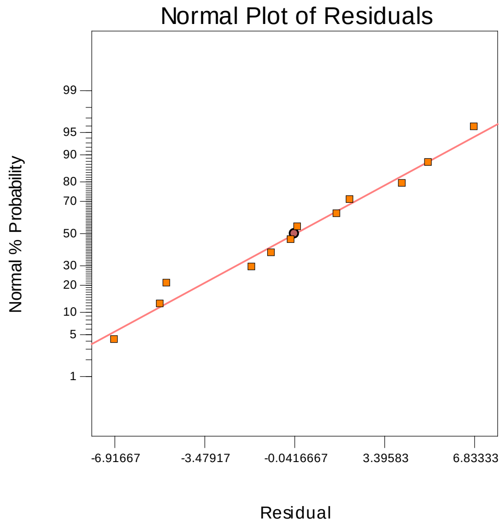
The model shown in the *Design Expert* output below does not differ substantially from the model in Problem 6.1, part (c).

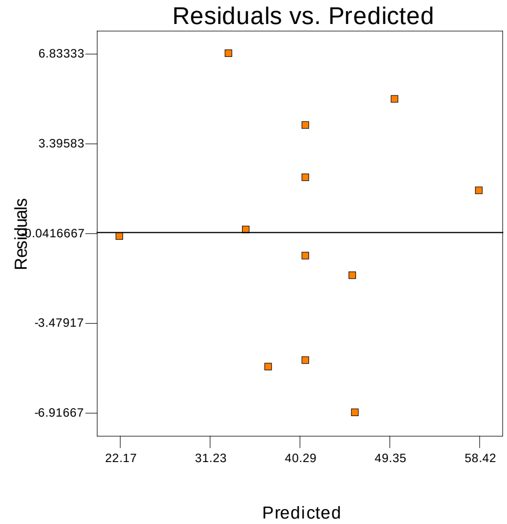
Design Expert Output

Final Equation in Terms of Coded Factors:

Life	=
+40.88	
+0.62	* A
+6.37	* B
+4.87	* C
-6.88	* A * C

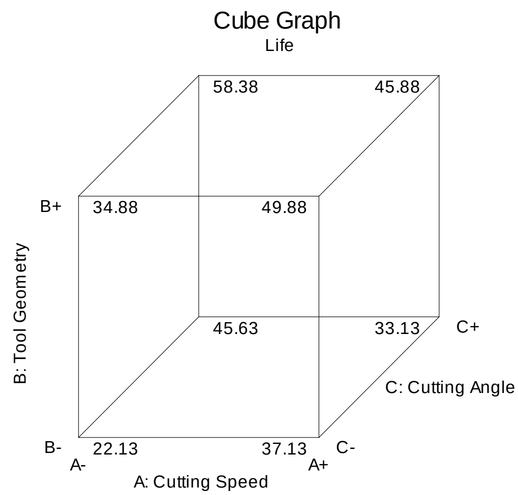
(d) Analyze the residuals.





(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	*d*	98	95
a	74	78	*ad*	72	76
b	81	85	*bd*	87	83

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

(a) Estimate the factor effects.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	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:

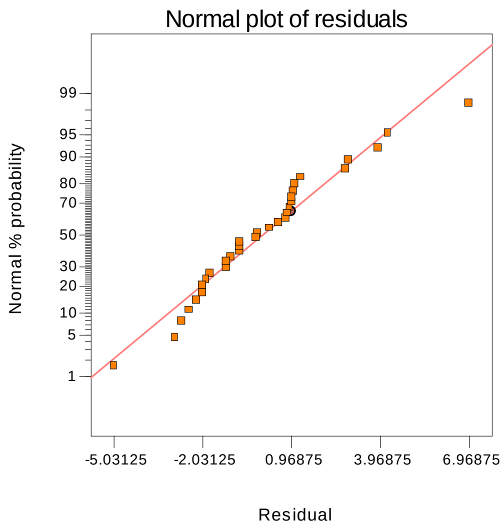
Design Expert Output

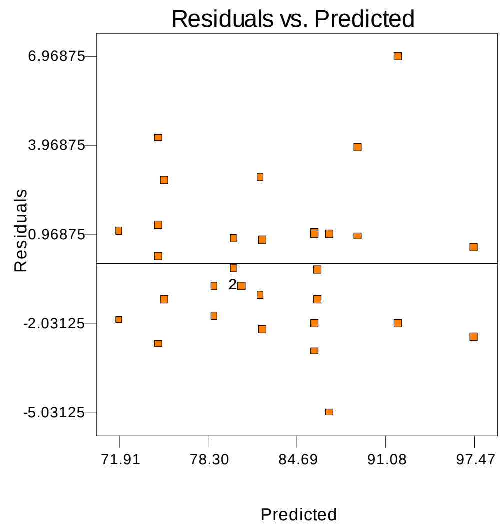
```
Final Equation in Terms of Coded Factors:

          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.

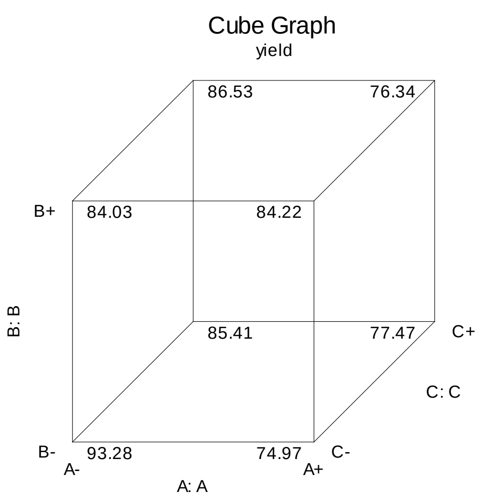
(d) Plot the residuals versus the predicted yield and on a normal probability scale. Does the residual analysis appear satisfactory?

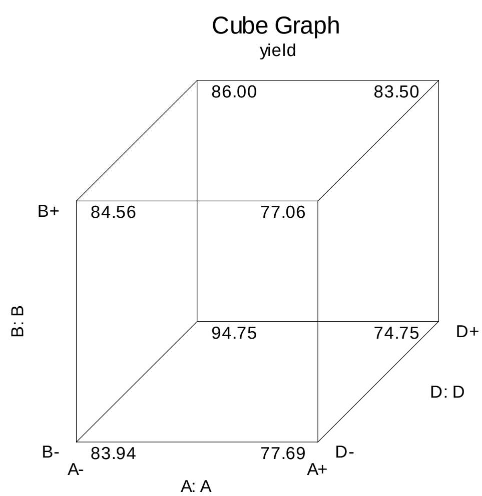
There appears to be one large residual both in the normal probability plot and in the plot of residuals versus predicted.





(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.

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

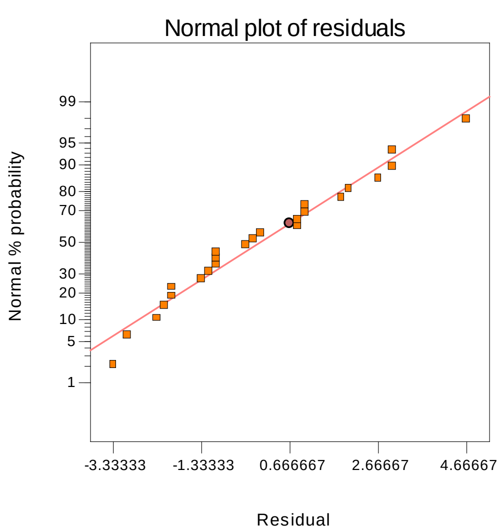
Design Expert Output

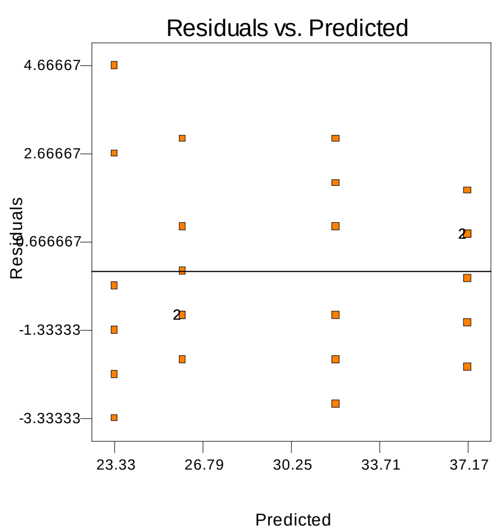
Response: Virus growth
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

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				

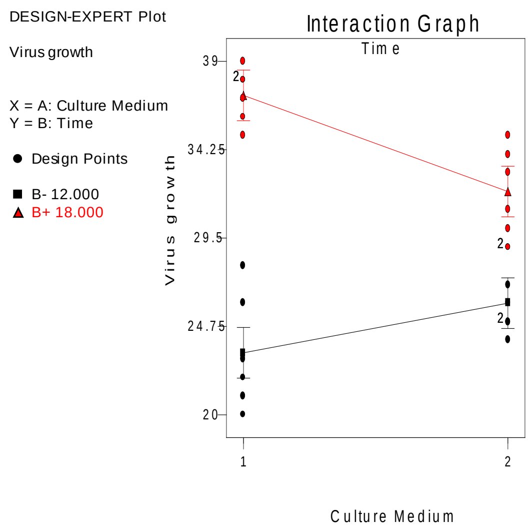
The Model F-value of 45.12 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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.

	Worker			
Bottle Type	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

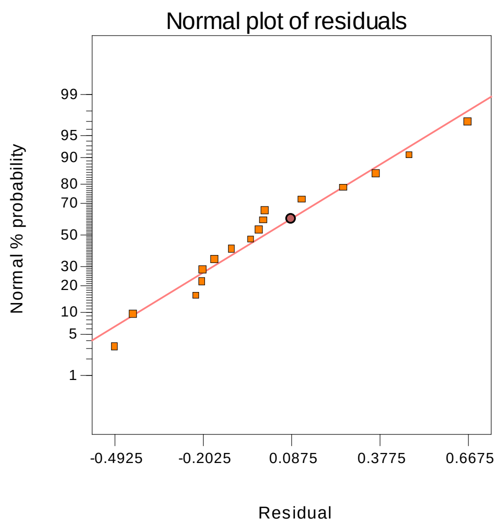
Response:Times
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

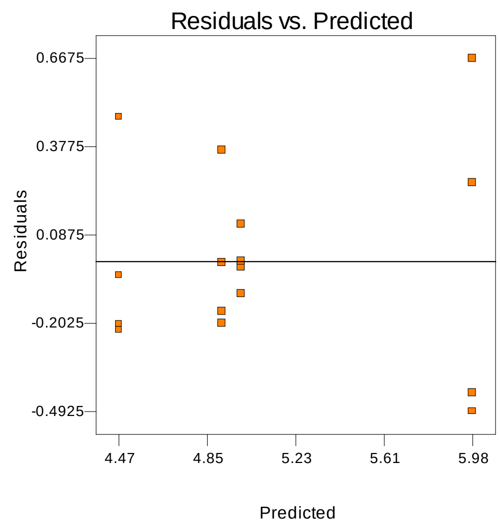
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	4.86	3	1.62	13.04	0.0004	significant
A	*2.02*	*1*	*2.02*	*16.28*	*0.0017*	
B	*2.54*	*1*	*2.54*	*20.41*	*0.0007*	
AB	*0.30*	*1*	*0.30*	*2.41*	*0.1463*	
Residual	1.49	12	0.12			
Lack of Fit	*0.000*	*0*				
Pure Error	*1.49*	*12*	*0.12*			
Cor Total	6.35	15				

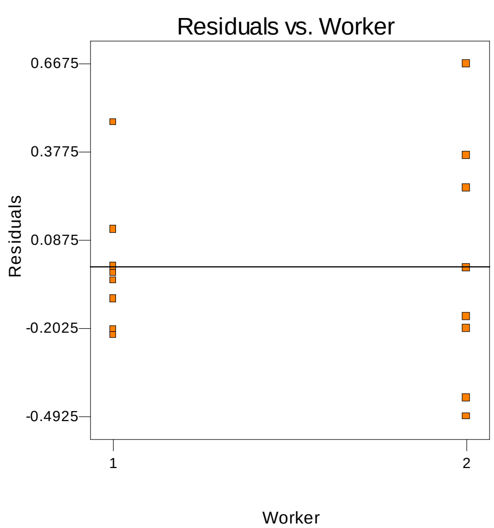
The Model F-value of 13.04 implies the model is significant. There is only a 0.04% chance that a "Model F-Value" this large could occur due to noise.

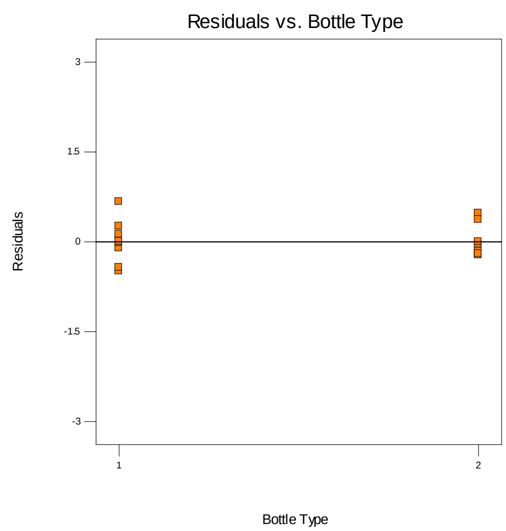
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B are significant model terms.

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.

	Worker			
Bottle Type	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

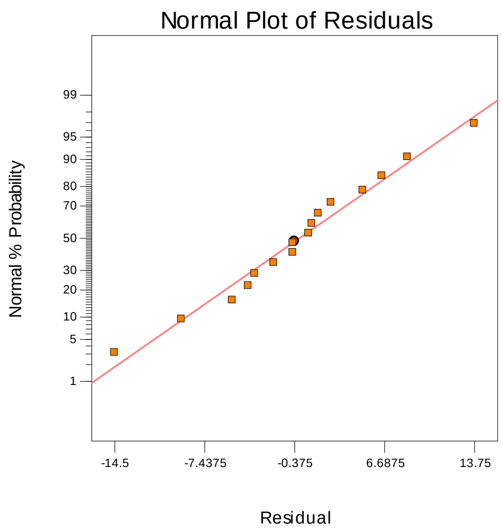
ANOVA for Selected Factorial Model

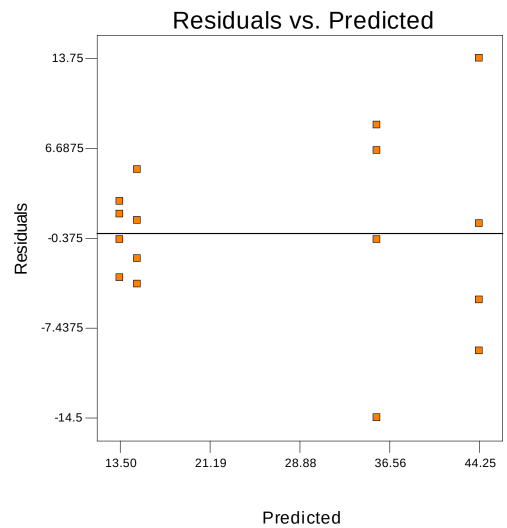
Analysis of variance table [Partial sum of squares]

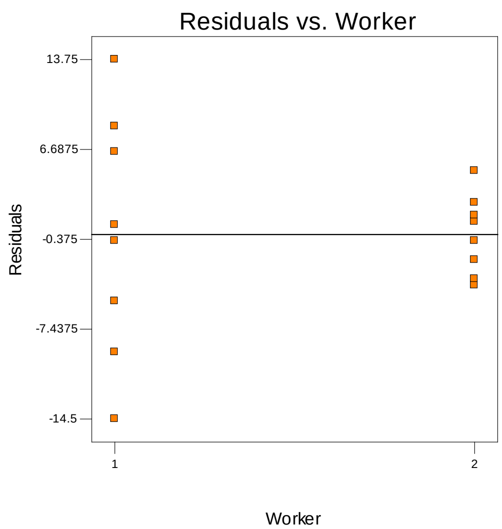
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	2784.19	3	928.06	16.03	0.0002	significant
A	*2626.56*	*1*	*2626.56*	*45.37*	*< 0.0001*	
B	*105.06*	*1*	*105.06*	*1.81*	*0.2028*	
AB	*52.56*	*1*	*52.56*	*0.91*	*0.3595*	
Residual	694.75	12	57.90			
Lack of Fit	*0.000*	*0*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

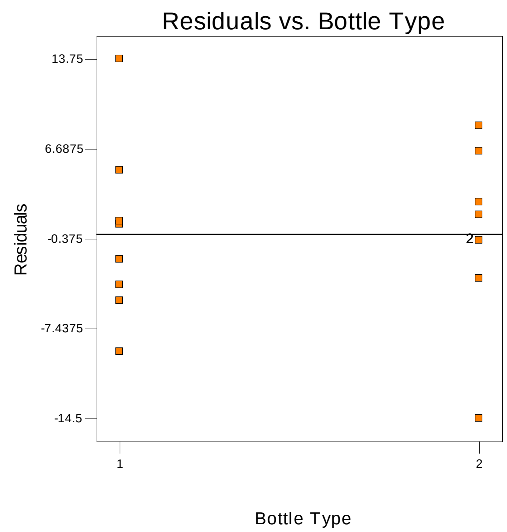
The Model F-value of 16.03 implies the model is significant. There is only a 0.02% chance that a "Model F-Value" this large could occur due to noise.

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	±3.80(1.96)= ±7.448
B	-5.125	±3.80(1.96)= ±7.448
AB	3.625	±3.80(1.96)= ±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:

		Replicate					Factor	Levels
A	*B*	I	II	III	IV		Low (-)	High (+)
-	-	14.037	16.165	13.972	13.907	*A*	55%	59%
+	-	13.880	13.860	14.032	13.914			
-	+	14.821	14.757	14.843	14.878	*B*	Short	Long
+	+	14.888	14.921	14.415	14.932		(10 min)	(15 min)

(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

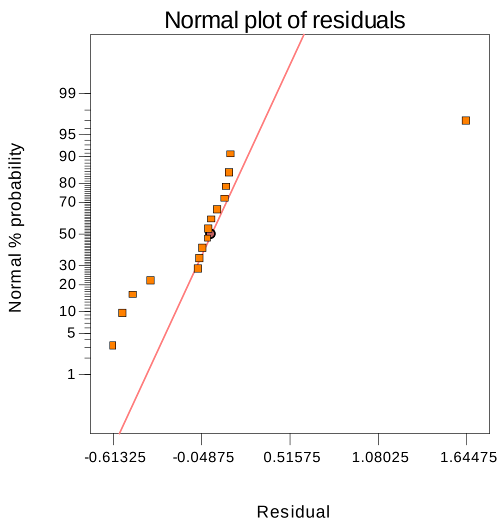
Final Equation in Terms of Coded Factors:

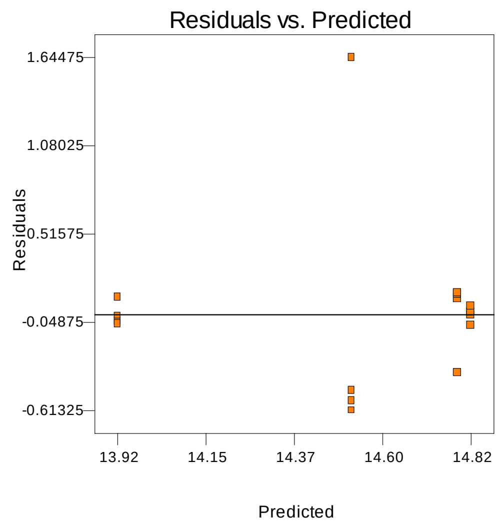
Thickness =
+14.51
-0.16 * A
+0.29 * B
+0.14 * A * B

Final Equation in Terms of Actual Factors:

Thickness =
+37.62656
-0.43119 * Flow Rate
-1.48735 * Dep Time
+0.028150 * Flow Rate * Dep Time

(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.

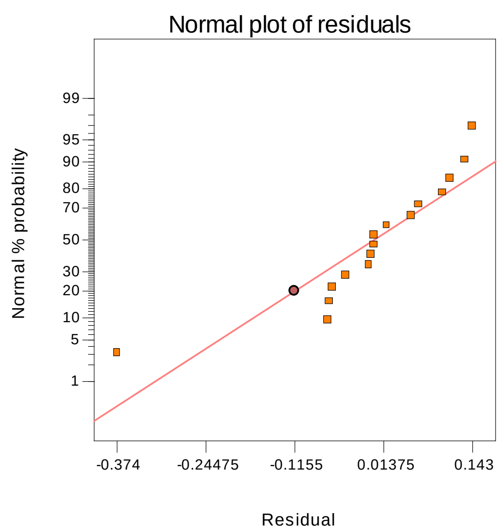
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B are significant model terms.

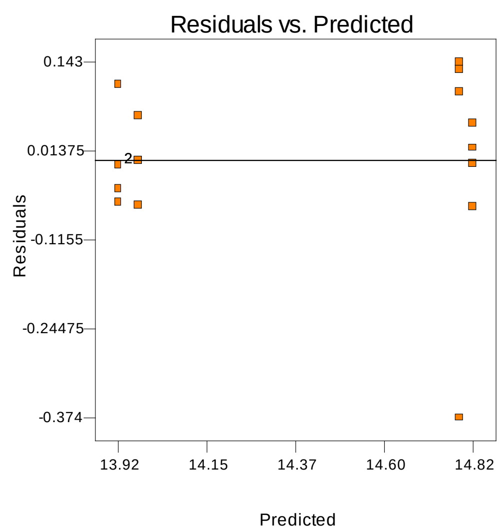
Final Equation in Terms of Coded Factors:

Thickness	=
+14.38	
-0.022	* A
+0.43	* B
+3.688E-003	* A * B

Final Equation in Terms of Actual Factors:

Thickness	=
+13.36650	
-0.020000	* Flow Rate
+0.12999	* Dep Time
+7.37500E-004	* Flow Rate * 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.

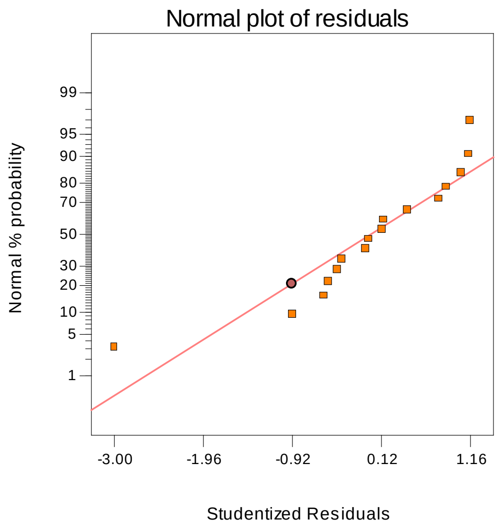
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case B are significant model terms.

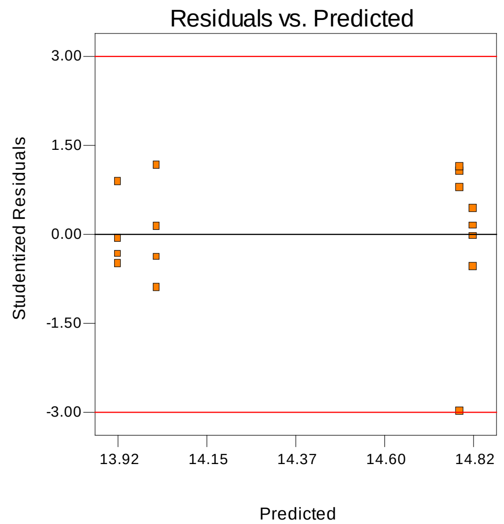
Final Equation in Terms of Coded Factors:

Thickness	=
+14.39	
-0.034	* A
+0.42	* B
+0.016	* A * B

Final Equation in Terms of Actual Factors:

Thickness	=
+15.50156	
-0.056188	* Flow Rate
-0.012350	* Dep Time
+3.15000E-003	* Flow Rate * Dep Time

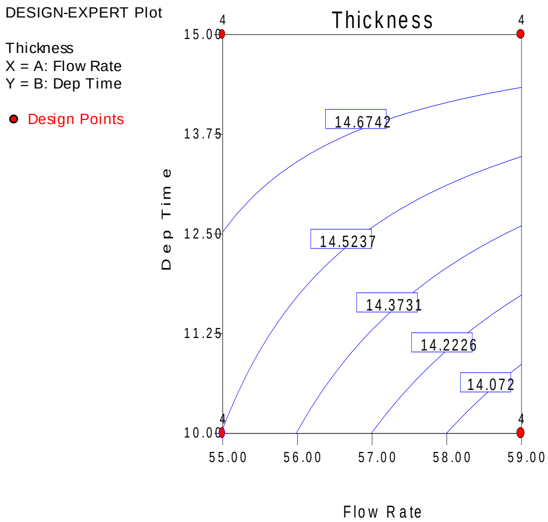


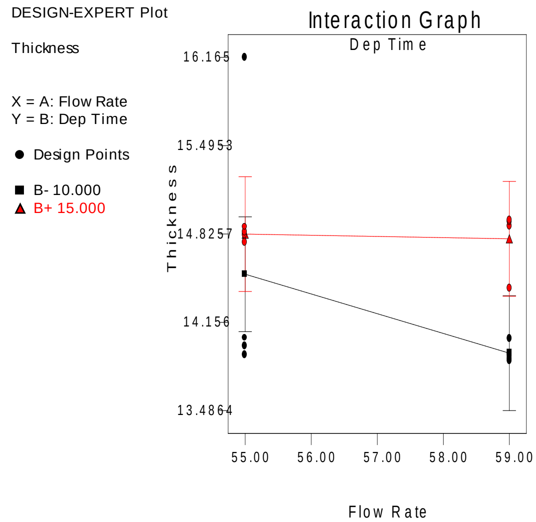


Another outlier is present and should be investigated.

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

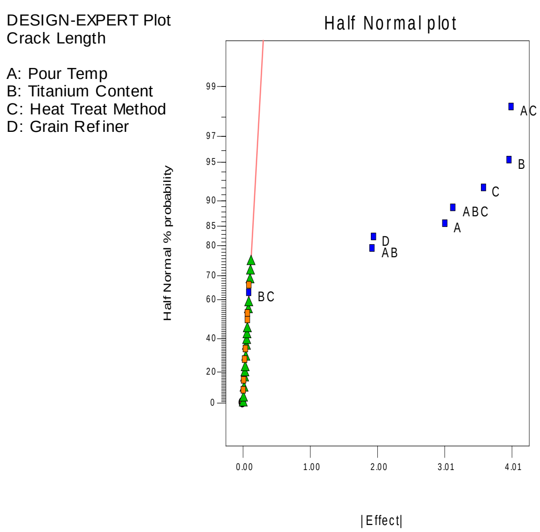
+	-	-	+	*ad*	16.867	17.052
-	+	-	+	*bd*	13.876	13.658
+	+	-	+	*abd*	19.824	19.639
-	-	+	+	*cd*	11.846	12.337
+	-	+	+	*acd*	6.125	5.904
-	+	+	+	*bcd*	11.190	10.935
+	+	+	+	*abcd*	15.653	15.053

(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 α=0.05.

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

Design Expert Output

Response: Crack Lengthin 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.

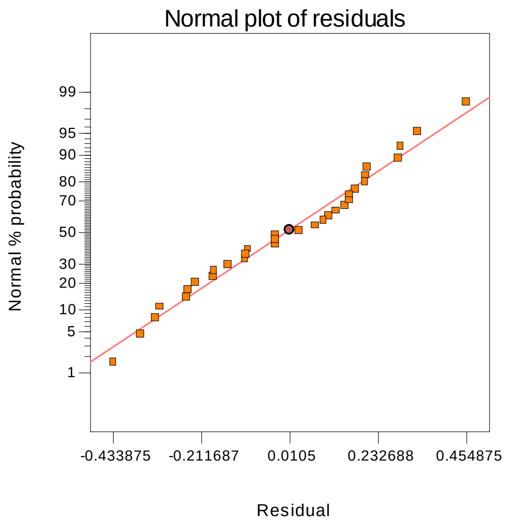
(c) Write down a regression model that can be used to predict crack length as a function of the significant main effects and interactions you have identified in part (b).

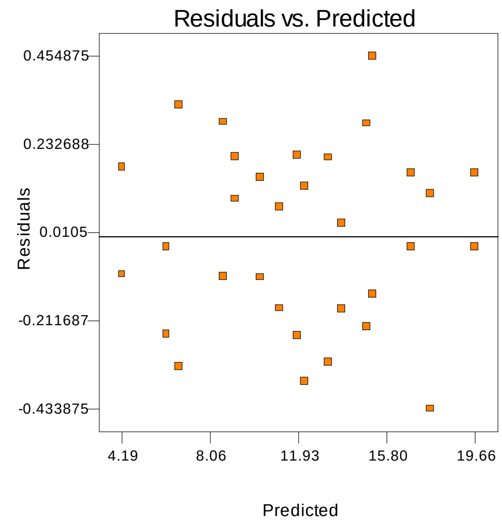
Design Expert Output

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.

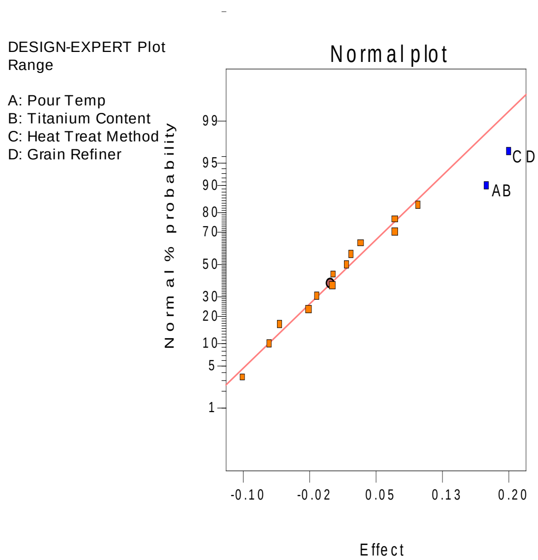




There is nothing unusual about the residuals.

(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	significant
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.

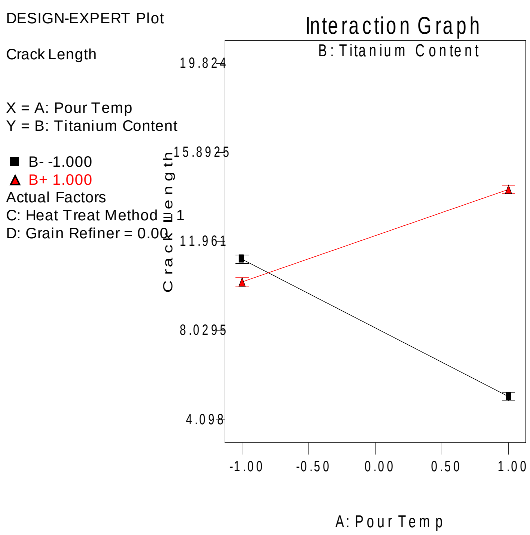
Final Equation in Terms of Coded Factors:

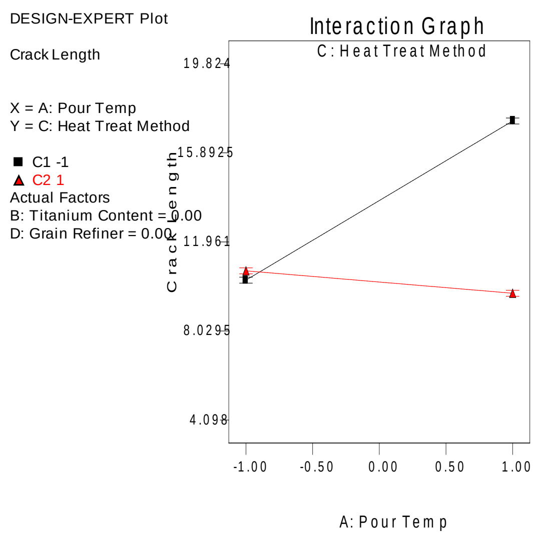
Range	=
+0.37	
+0.089	* A * B
+0.10	* C * D

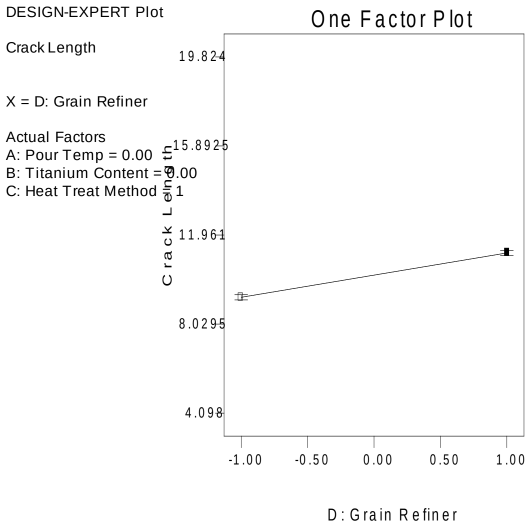
(f) What recommendations would you make regarding process operations? Use interaction and/or main effect plots to assist in drawing conclusions.

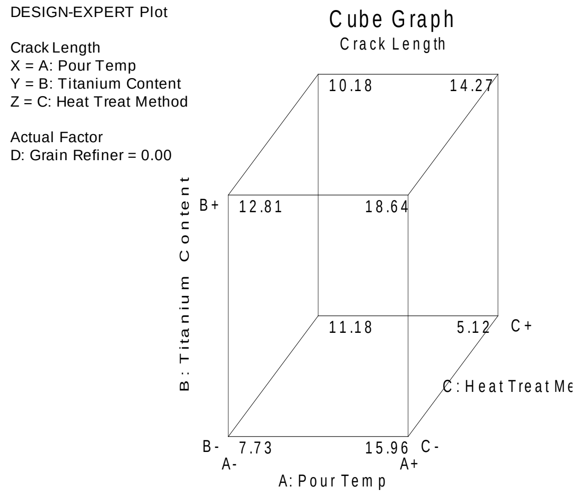
From the interaction plots, choose *A* at the high level and *B* at the low level. In each of these plots, *D* can be at either level. From the main effects plot of *C*, choose *C* at the high level. Based on the range analysis, with *C* at the high level, *D* should be set at the low level.

From the analysis of the crack length data:

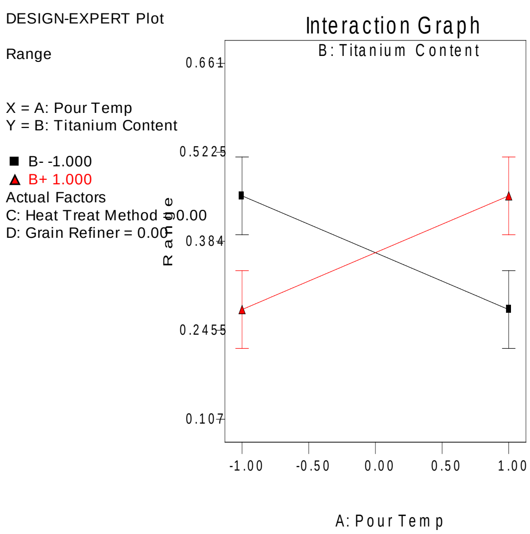


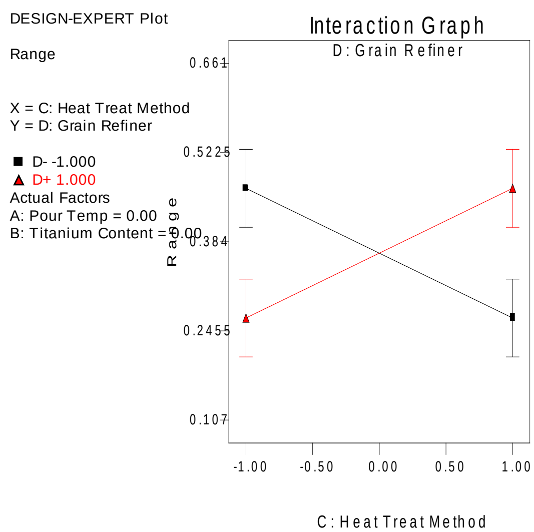






From the analysis of the ranges:





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.

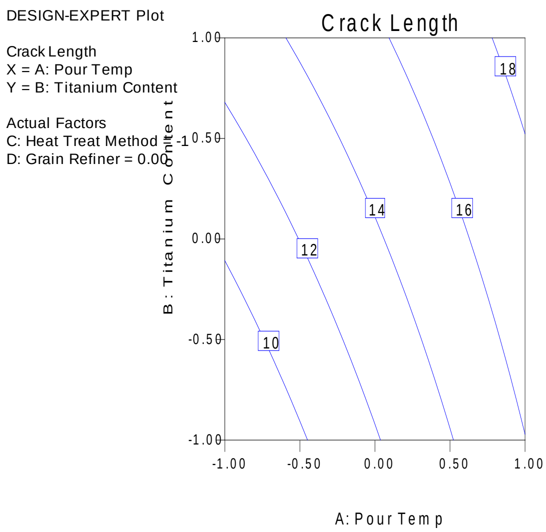
(a) Write two regression models for predicting crack length, one for each level of the heat treatment method variable. What differences, if any, do you notice in these two equations?

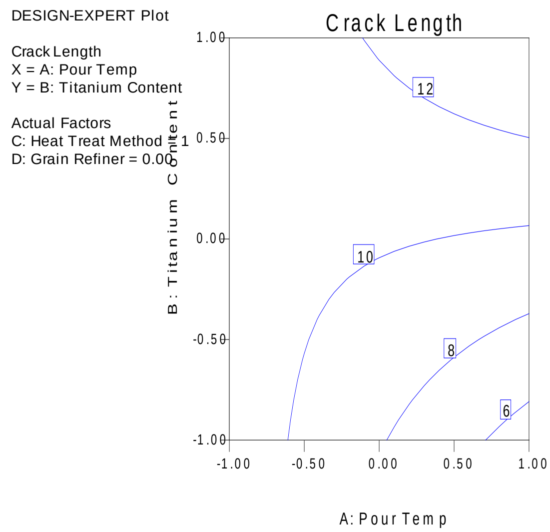
Design Expert Output

```
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*=-.

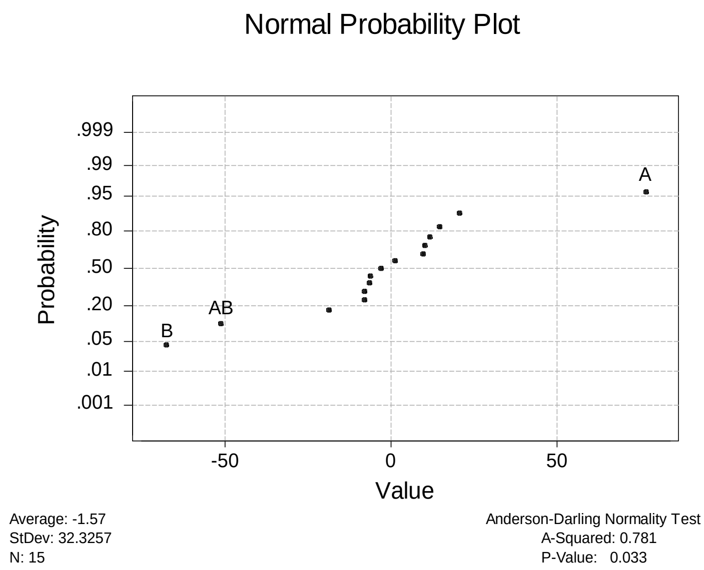
Low level of *A*, low level of *B*, and low level of *D*.

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.

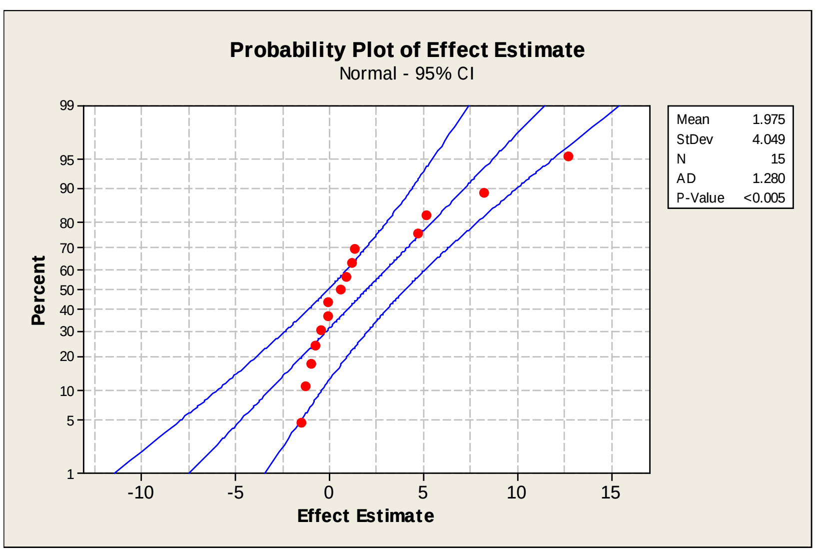
The plot from Minitab follows.



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

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

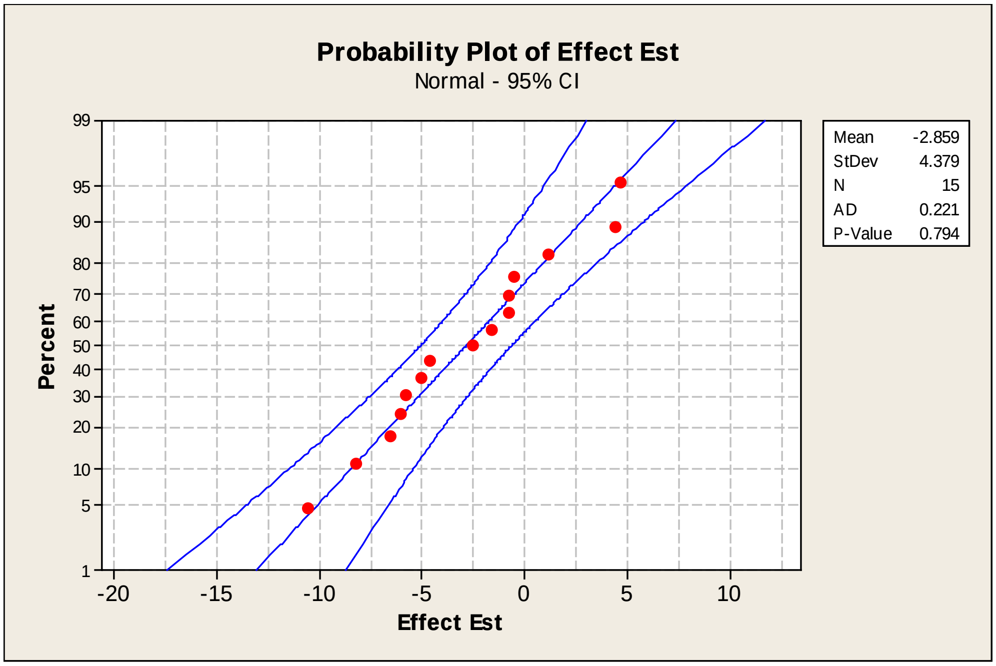
6.18. The effect estimates from a 2^4 factorial design are as follows: ABCD = -1.5138, ABC = -1.2661, ABD = -0.9852, ACD = -0.7566, BCD = -0.4842, CD = -0.0795, BD = -0.0793, AD = 0.5988, BC = 0.9216, AC = 1.1616, AB = 1.3266, D = 4.6744, C = 5.1458, B = 8.2469, and A = 12.7151. Are you comfortable with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

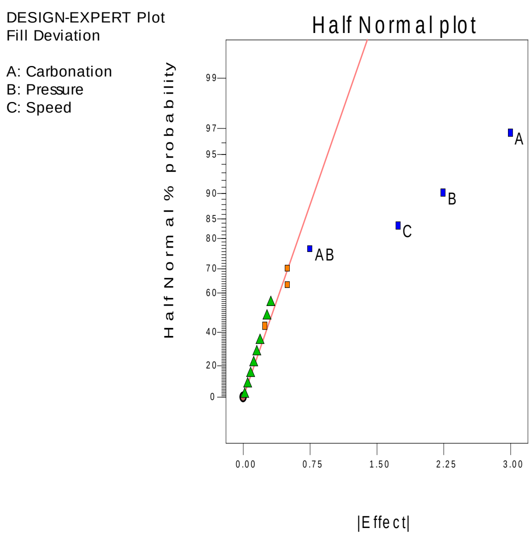
	Coded Factors			Fill Height Deviation	
Run	*A*	*B*	*C*	Replicate 1	Replicate 2
1	-	-	-	-3	-1
2	+	-	-	0	1
3	-	+	-	-1	0
4	+	+	-	2	3
5	-	-	+	-1	0
6	+	-	+	2	1
7	-	+	+	1	1
8	+	+	+	6	5

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

B (psi)	25	30
C (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
A	*36.00*	*1*	*36.00*	*54.62*	*< 0.0001*	
B	*20.25*	*1*	*20.25*	*30.72*	*0.0002*	
C	*12.25*	*1*	*12.25*	*18.59*	*0.0012*	
AB	*2.25*	*1*	*2.25*	*3.41*	*0.0917*	
Residual	7.25	11	0.66			
Lack of Fit	*2.25*	*3*	*0.75*	*1.20*	*0.3700*	*not significant*
Pure Error	*5.00*	*8*	*0.63*			
Cor Total	78.00	15				

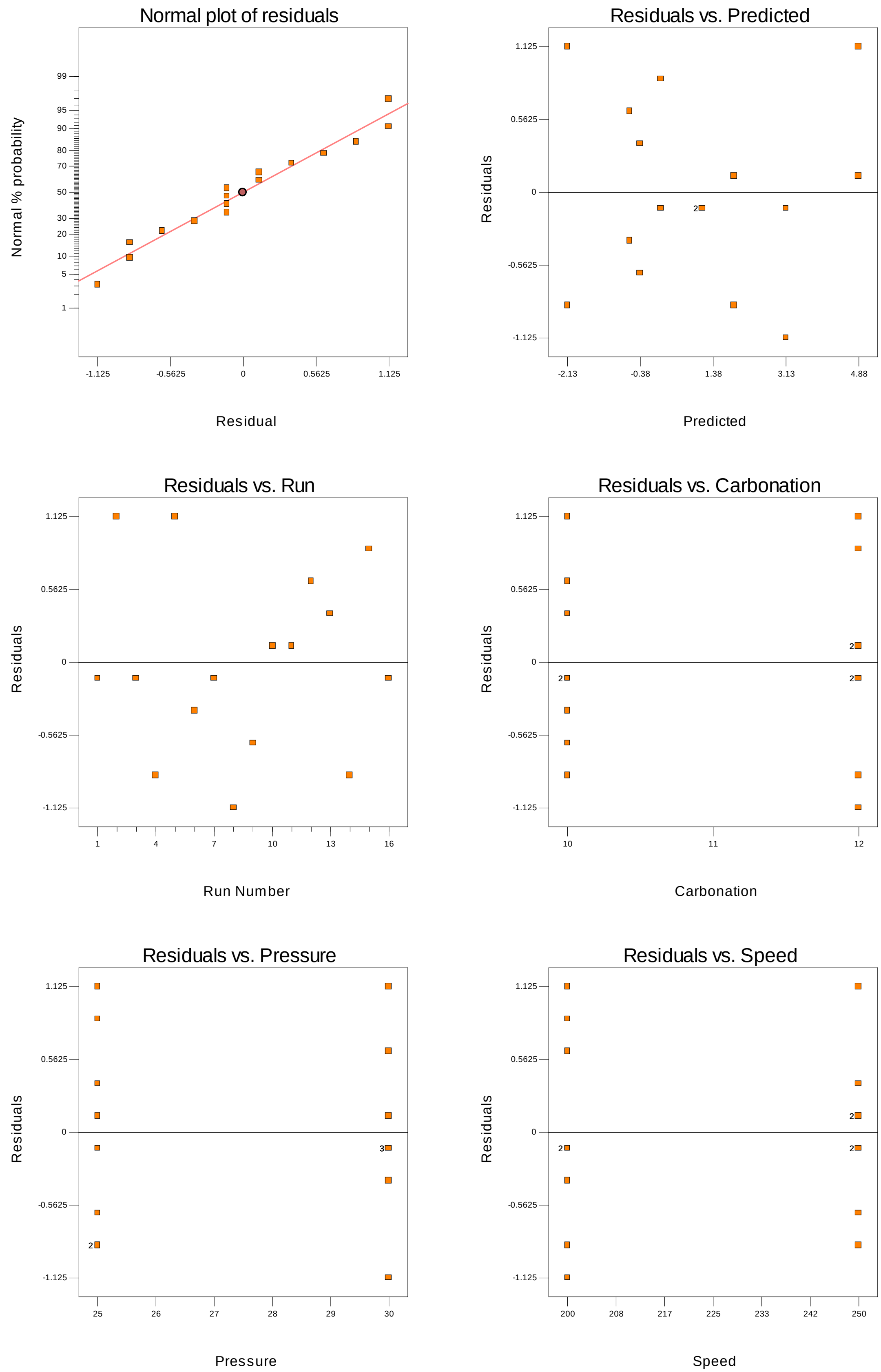
The Model F-value of 26.84 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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



Normal plot of residuals
Normal % probability
Residual
Residuals vs. Predicted
Residuals
Predicted
Residuals vs. Run
Run Number
Residuals vs. Carbonation
Carbonation
Residuals vs. Pressure
Pressure
Residuals vs. Speed
Speed

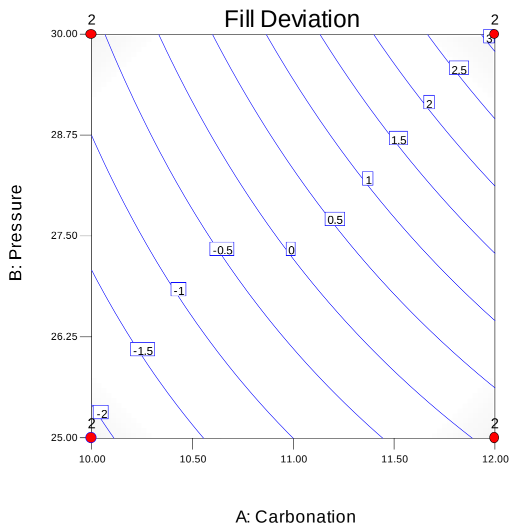
(c) Obtain a model for predicting fill height deviation in terms of the important process variables. Use this model to construct contour plots to assist in interpreting the results of the experiment.

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

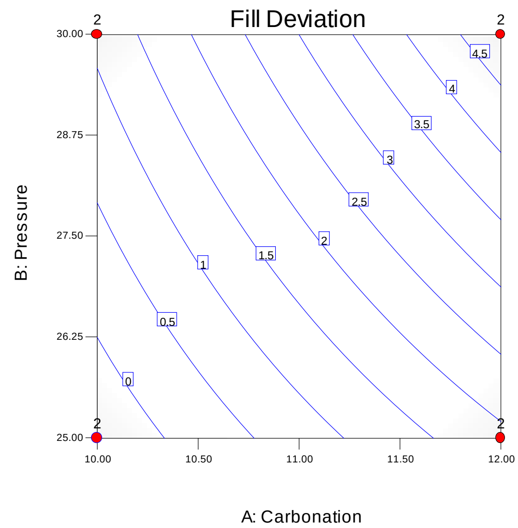
Design Expert Output

```
Final Equation in Terms of Coded Factors

            Fill Deviation  =
                     +1.00
                     +1.50   * A
                     +1.13   * B
                     +0.88   * C
                     +0.38   * A * B

Final Equation in Terms of Actual Factors

            Fill Deviation  =
                  +9.62500
                  -2.62500   * Carbonation
                  -1.20000   * Pressure
                 +0.035000   * Speed
                  +0.15000   * Carbonation * Pressure
```

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.

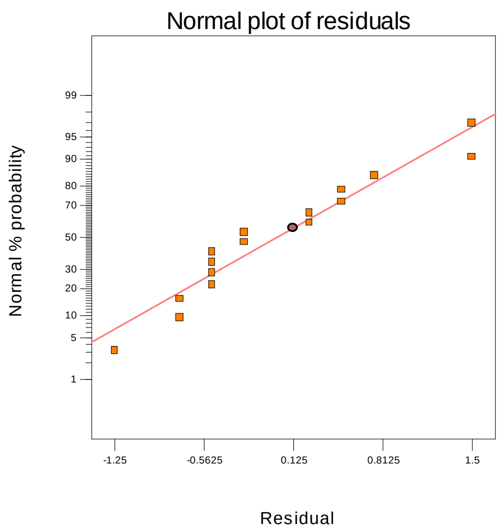
Std. Dev.	0.89	R-Squared	0.8782
Mean	1.00	Adj R-Squared	0.8478
C.V.	88.98	Pred R-Squared	0.7835
PRESS	16.89	Adeq Precision	15.735

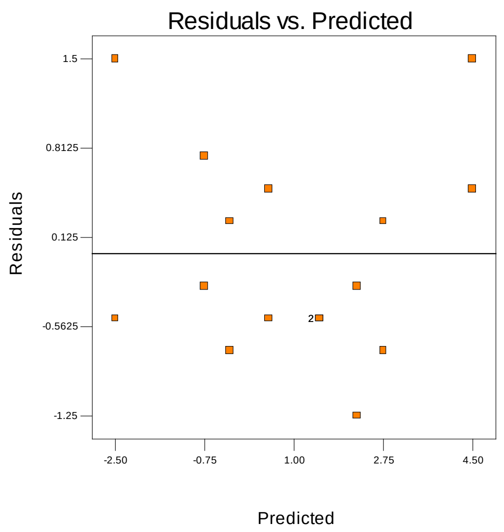
Final Equation in Terms of Coded Factors

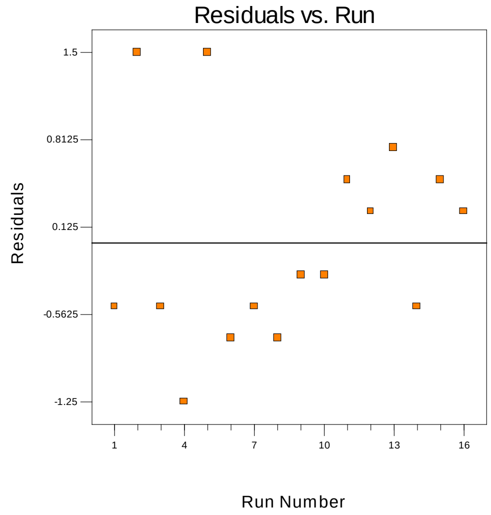
Fill Deviation	=
+1.00	
+1.50	* A
+1.13	* B
+0.88	* C

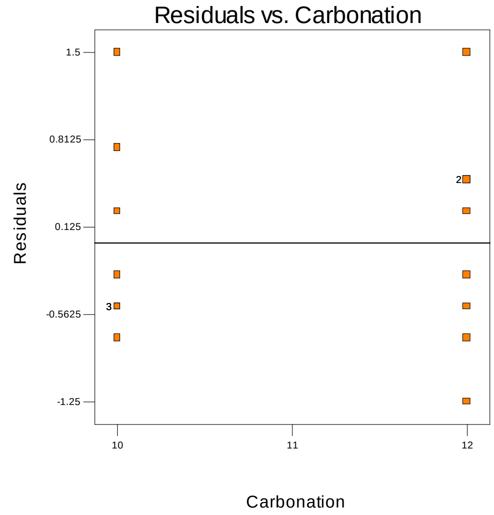
Final Equation in Terms of Actual Factors

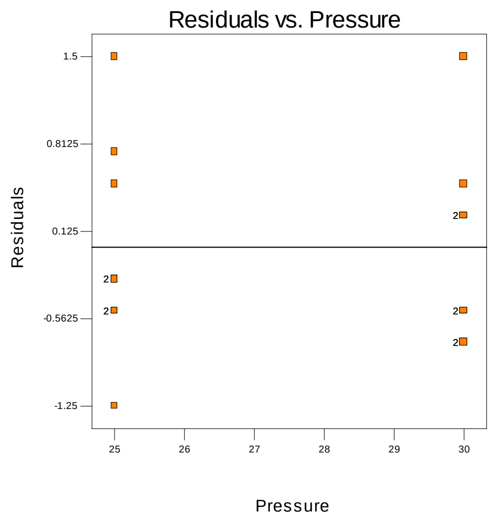
Fill Deviation	=
-35.75000	
+1.50000	* Carbonation
+0.45000	* Pressure
+0.035000	* Speed

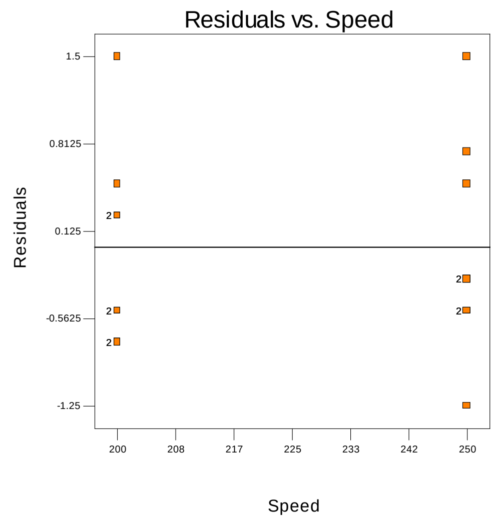


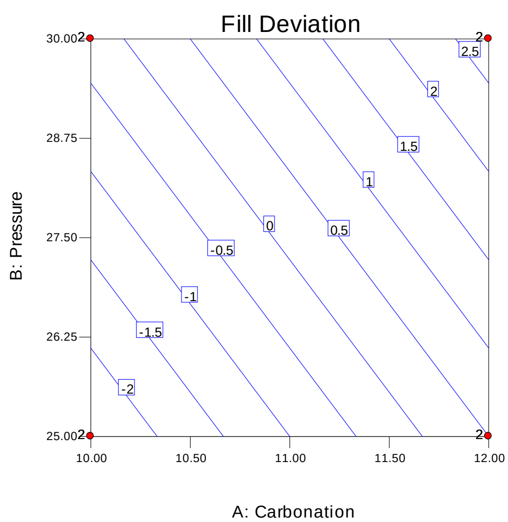




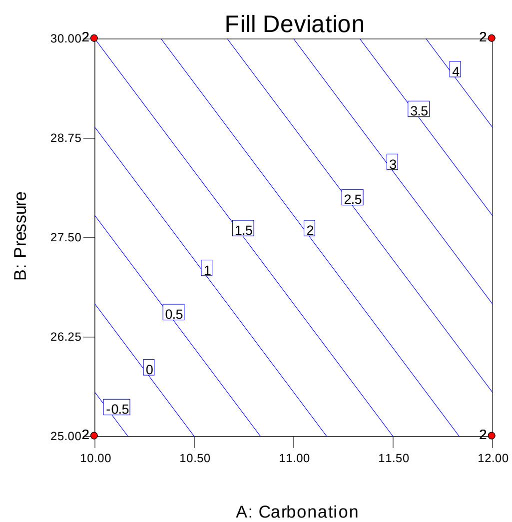








Speed set at 200 b/m



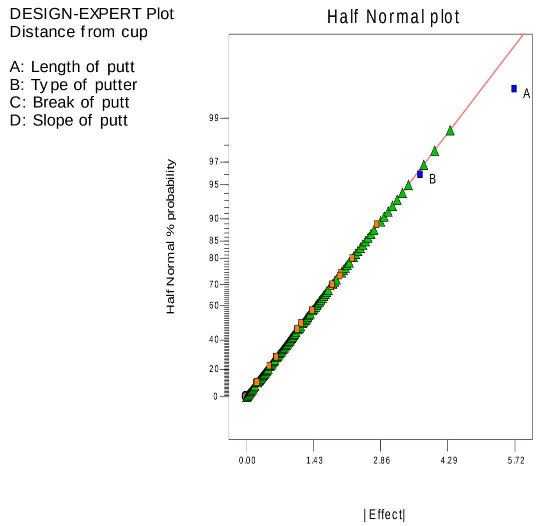
Speed set at 250 b/m

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				

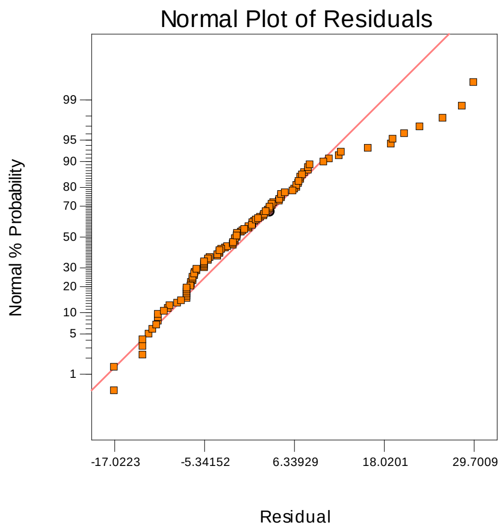
The Model F-value of 7.69 implies the model is significant. There is only a 0.08% chance that a "Model F-Value" this large could occur due to noise.

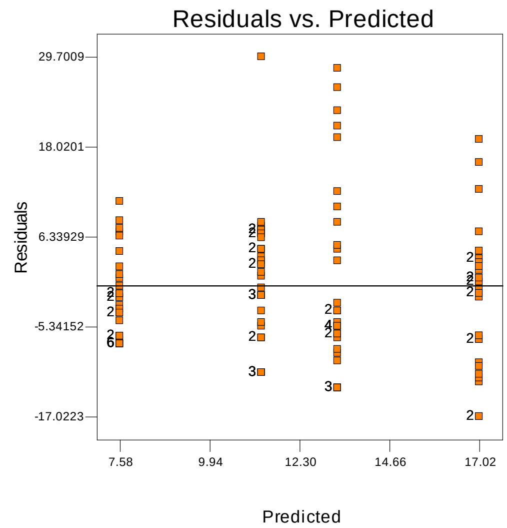
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, are significant model terms.

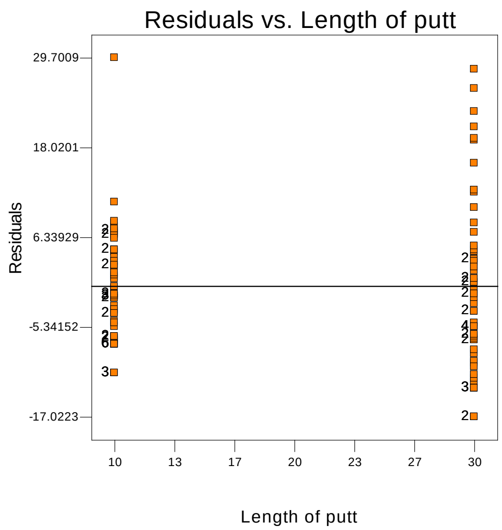
Std. Dev.	9.21	R-Squared	0.1237
Mean	12.30	Adj R-Squared	0.1076
C.V.	74.90	Pred R-Squared	0.0748
PRESS	9765.06	Adeq Precision	6.266

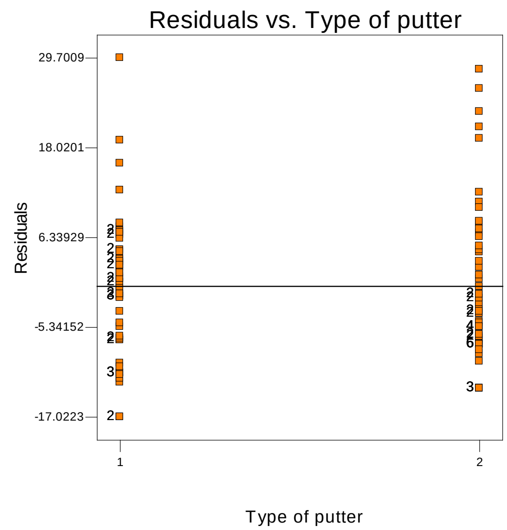
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

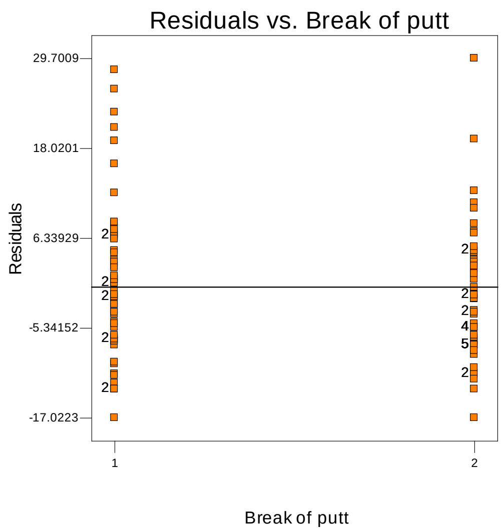
The residual plots for the model containing only the significant factors A and B are shown below. The normality assumption appears to be violated. Also, as a golfer might expect, there is a slight inequality of variance with regards to the length of putt. A square root transformation is applied which 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.

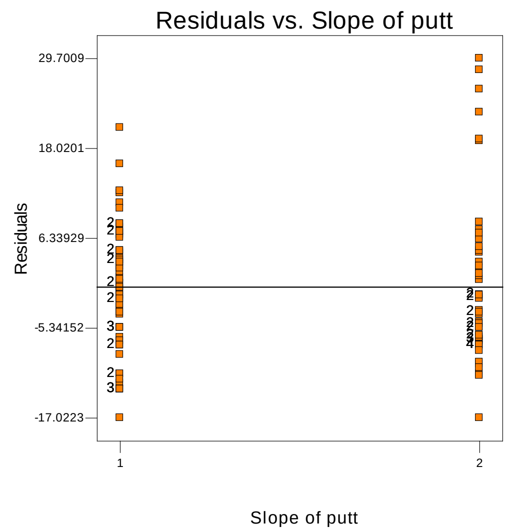

Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-17.0223
-5.34152
6.33929
18.0201
29.7009
Residual


Residuals vs. Predicted
Residuals
29.7009
18.0201
6.33929
-5.34152
-17.0223
7.58
9.94
12.30
14.66
17.02
Predicted


Residuals vs. Length of putt
Residuals
29.7009
18.0201
6.33929
-5.34152
-17.0223
10
13
17
20
23
27
30
Length of putt


Residuals vs. Type of putter
Residuals
29.7009
18.0201
6.33929
-5.34152
-17.0223
1
2
Type of putter


Residuals vs. Break of putt
Residuals
29.7009
18.0201
6.33929
-5.34152
-17.0223
1
2
Break of putt


Residuals vs. Slope of putt
Residuals
29.7009
18.0201
6.33929
-5.34152
-17.0223
1
2
Slope of putt

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

Response: Distance from cupTransform:Square root Constant: 0

ANOVA for Selected Factorial Model

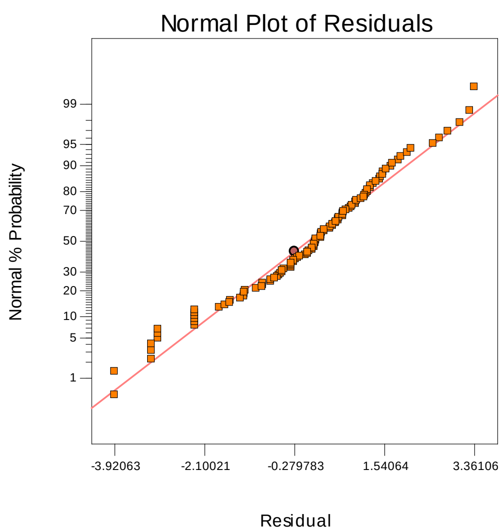
Analysis of variance table [Terms added sequentially (first to last)]

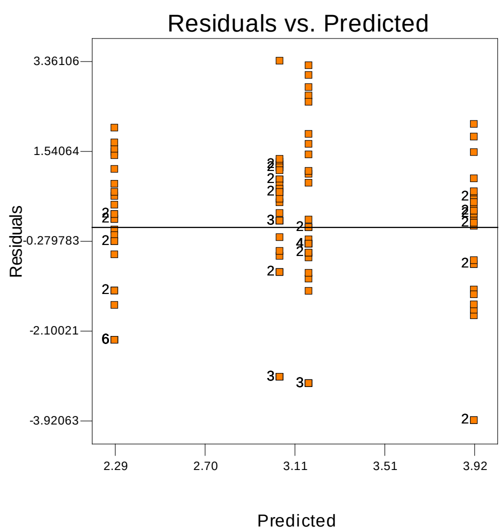
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	37.26	2	18.63	7.85	0.0007	significant
A	*21.61*	*1*	*21.61*	*9.11*	*0.0032*	
B	*15.64*	*1*	*15.64*	*6.59*	*0.0116*	
Residual	258.63	109	2.37			
Lack of Fit	*30.19*	*13*	*2.32*	*0.98*	*0.4807*	*not significant*
Pure Error	*228.45*	*96*	*2.38*			
Cor Total	295.89	111				

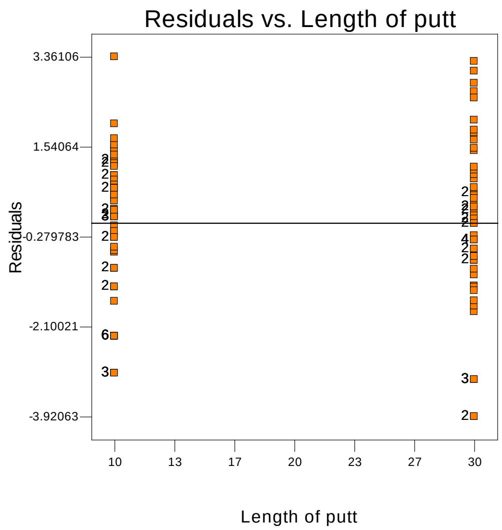
The Model F-value of 7.85 implies the model is significant. There is only a 0.07% chance that a "Model F-Value" this large could occur due to noise.

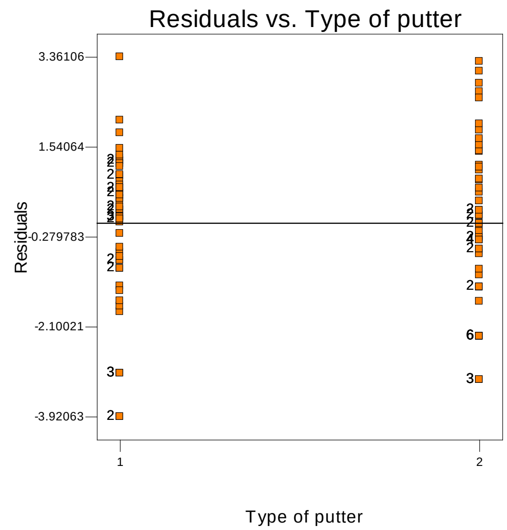
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, are significant model terms.

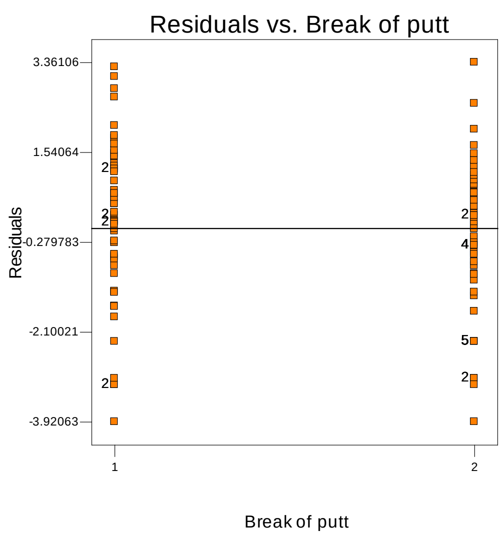
Std. Dev.	1.54	R-Squared	0.1259
Mean	3.11	Adj R-Squared	0.1099
C.V.	49.57	Pred R-Squared	0.0771
PRESS	273.06	Adeq Precision	6.450

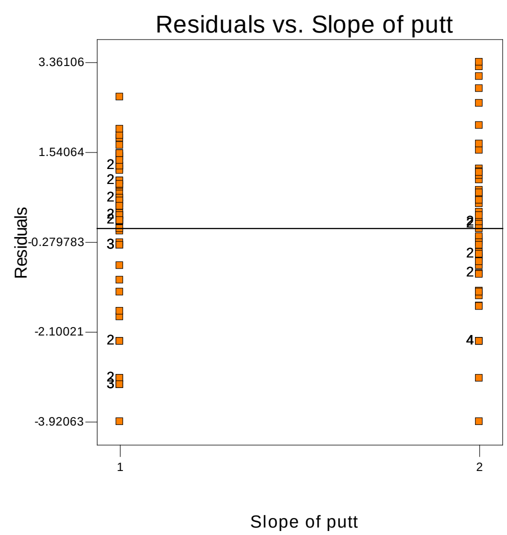


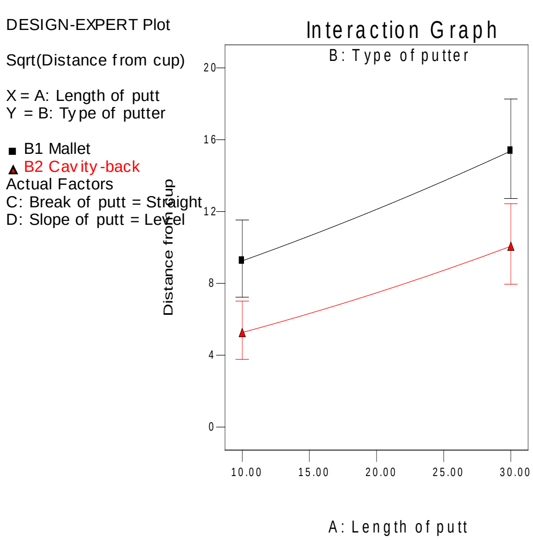



Residuals vs. Length of putt
Residuals
3.36106
1.54064
-0.279783
-2.10021
-3.92063
10
13
17
20
23
27
30
Length of putt


Residuals vs. Type of putter
Residuals
3.36106
1.54064
-0.279783
-2.10021
-3.92063
1
2
Type of putter


Residuals vs. Break of putt
Residuals
3.36106
1.54064
-0.279783
-2.10021
-3.92063
1
2
Break of putt


Residuals vs. Slope of putt
Residuals
3.36106
1.54064
-0.279783
-2.10021
-3.92063
1
2
Slope of putt

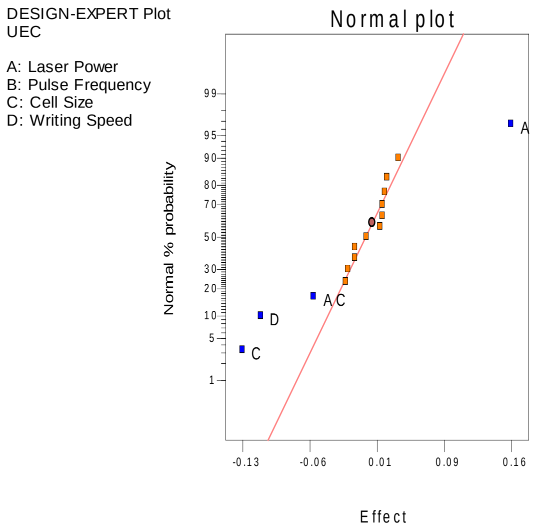

DESIGN-EXPERT Plot
Sqrt(Distance from cup)
X = A: Length of putt
Y = B: Type of putter
B1 Mallet
B2 Cavity-back
Actual Factors
C: Break of putt = Straight
D: Slope of putt = Level
Interaction Graph
B: Type of putter
Distance from cup
20
16
12
8
4
0
10.00
15.00
20.00
25.00
30.00
A: Length of putt

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.

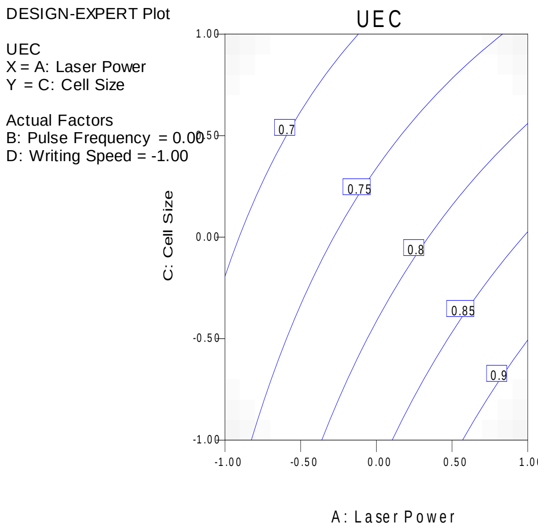
Std. Dev.	0.041	R-Squared	0.9281
Mean	0.72	Adj R-Squared	0.9020
C.V.	5.68	Pred R-Squared	0.8479
PRESS	0.039	Adeq Precision	17.799

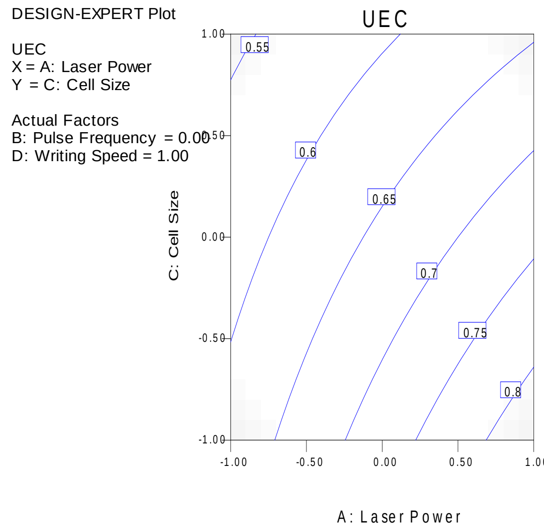
Final Equation in Terms of Coded Factors

UEC	=
+0.72	
+0.080	* A
-0.066	* C
-0.056	* D
-0.027	* A * C

Final Equation in Terms of Actual Factors

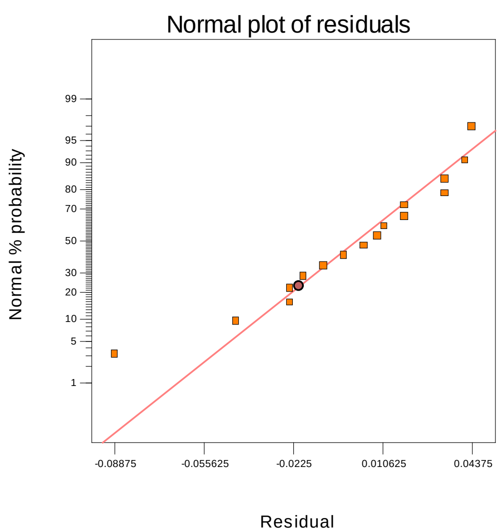
UEC	=
+0.71625	
+0.080000	* Laser Power
-0.066250	* Cell Size
-0.056250	* Writing Speed
-0.027500	* Laser Power * Cell Size

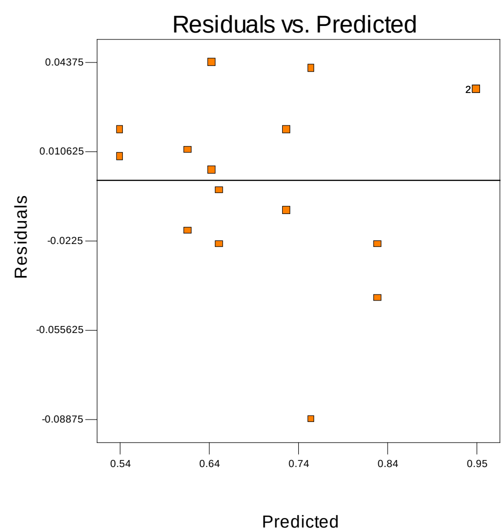


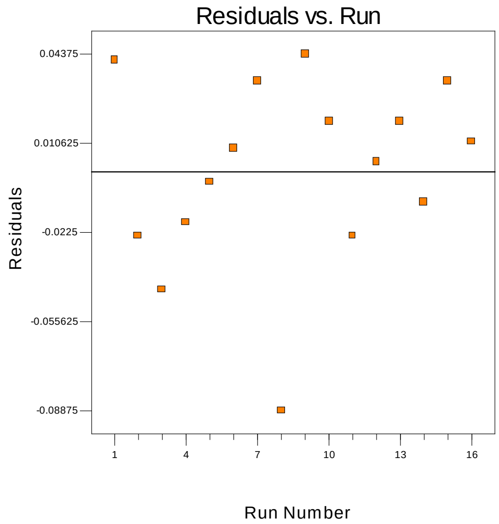


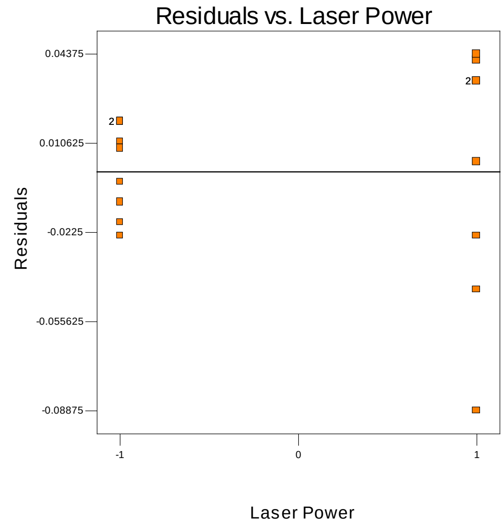
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy?

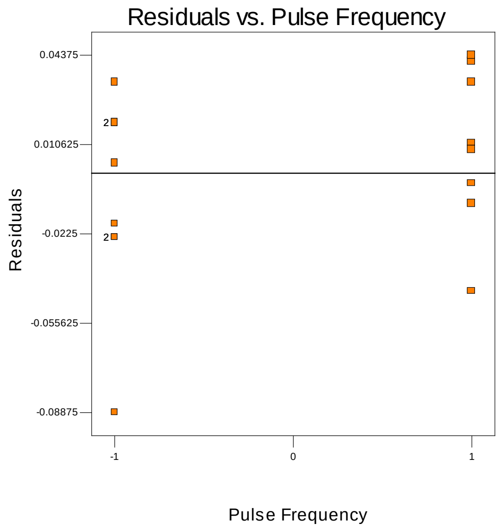
The residual plots appear acceptable with the exception of run 8, standard order 6. This value should be verified by the engineer.

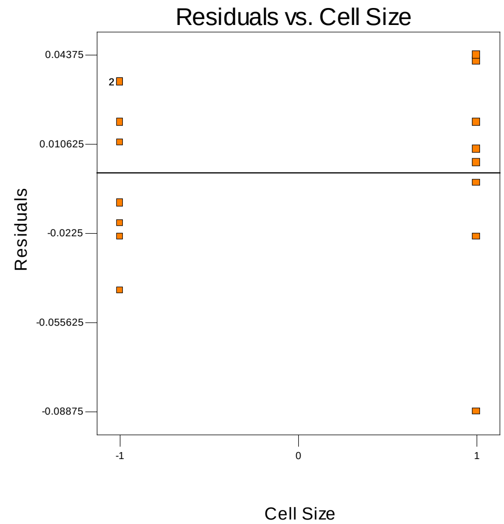


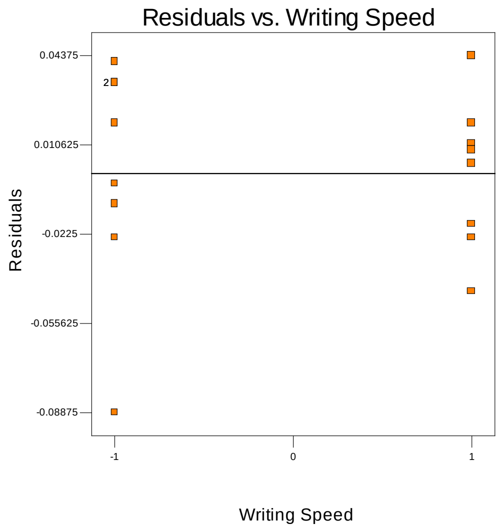



Residuals vs. Run
Residuals
0.04375
0.010625
-0.0225
-0.055625
-0.08875
1
4
7
10
13
16
Run Number


Residuals vs. Laser Power
Residuals
0.04375
0.010625
-0.0225
-0.055625
-0.08875
2
2
-1
0
1
Laser Power


Residuals vs. Pulse Frequency
Residuals
0.04375
0.010625
-0.0225
-0.055625
-0.08875
2
2
-1
0
1
Pulse Frequency


Residuals vs. Cell Size
Residuals
0.04375
0.010625
-0.0225
-0.055625
-0.08875
2
-1
0
1
Cell Size


Residuals vs. Writing Speed
Residuals
0.04375
0.010625
-0.0225
-0.055625
-0.08875
2
-1
0
1
Writing Speed

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	**Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI 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

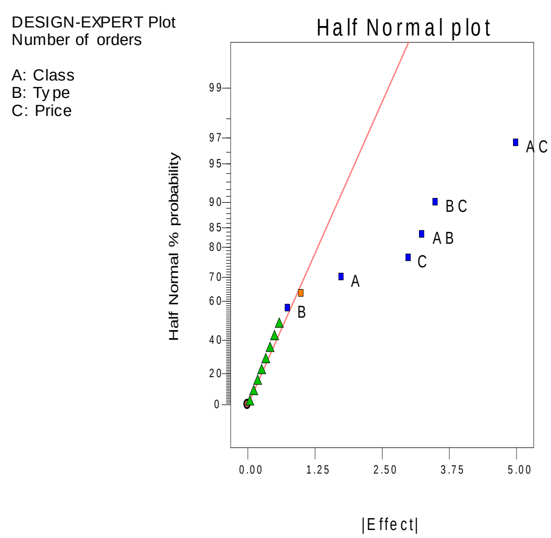
6.24. A company markets its products by direct mail. An experiment was conducted to study the effects of three factors on the customer response rate for a particular product. The three factors are *A* = type of mail used (3rd class, 1st class), *B* = type of descriptive brochure (color, black-and-white), 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.

	Coded Factors			Number of Orders	
Run	*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 hierarchal 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.

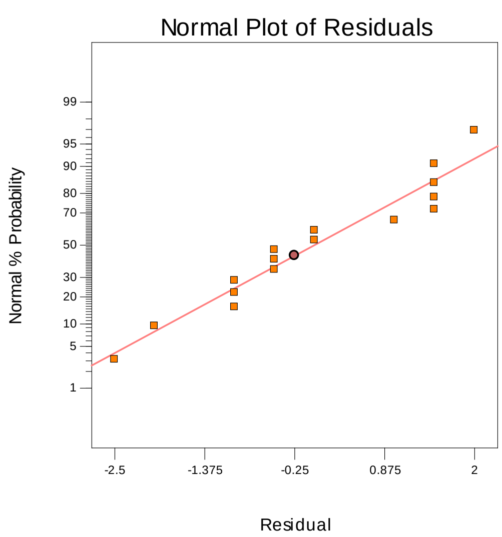
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case C, AB, AC, BC are significant model terms.

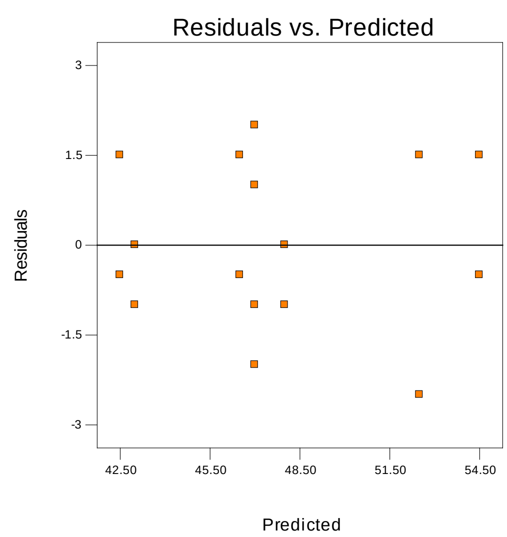
Std. Dev.	1.76	R-Squared	0.8962
Mean	47.63	Adj R-Squared	0.8270
C.V.	3.70	Pred R-Squared	0.6719
PRESS	88.49	Adeq Precision	10.286

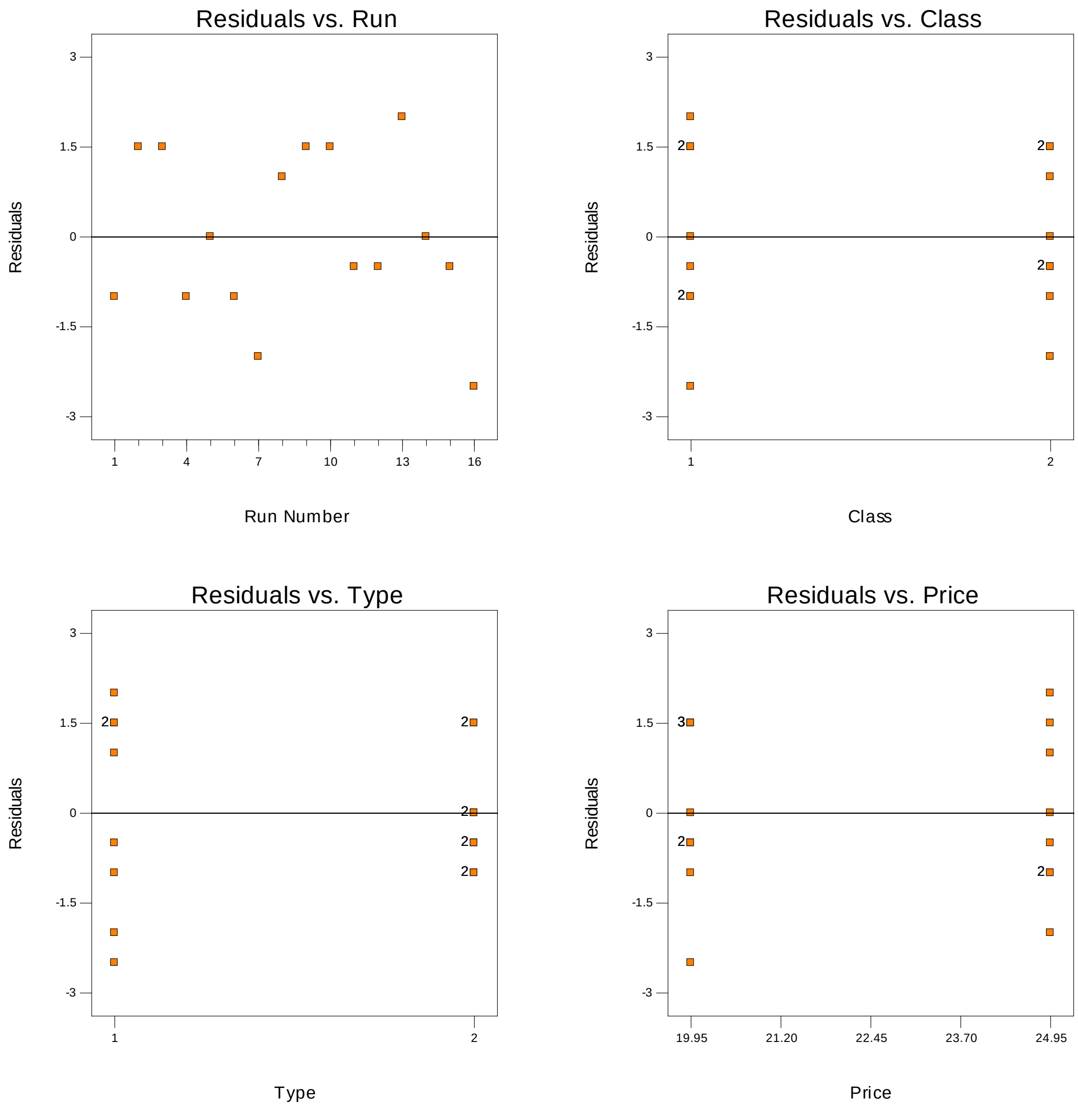
Coefficient Factor	**Standard Estimate**	**95% CI DF**	**95% CI Error**	**Low**	**High**	**VIF**
Intercept	47.63	1	0.44	46.63	48.62	
A-Class	-0.88	1	0.44	-1.87	0.12	1.00
B-Type	0.37	1	0.44	-0.62	1.37	1.00
C-Price	1.50	1	0.44	0.50	2.50	1.00
AB	1.63	1	0.44	0.63	2.62	1.00
AC	2.50	1	0.44	1.50	3.50	1.00
BC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

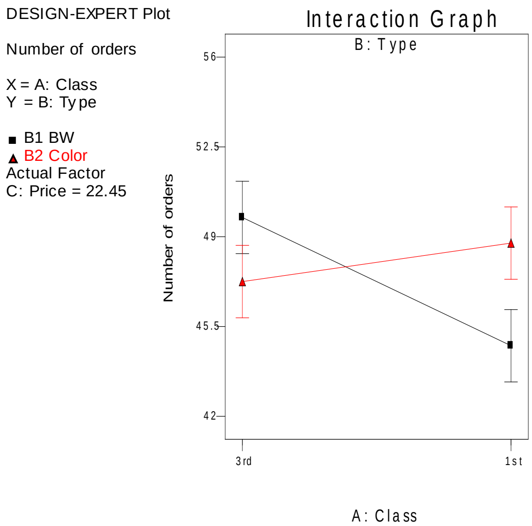


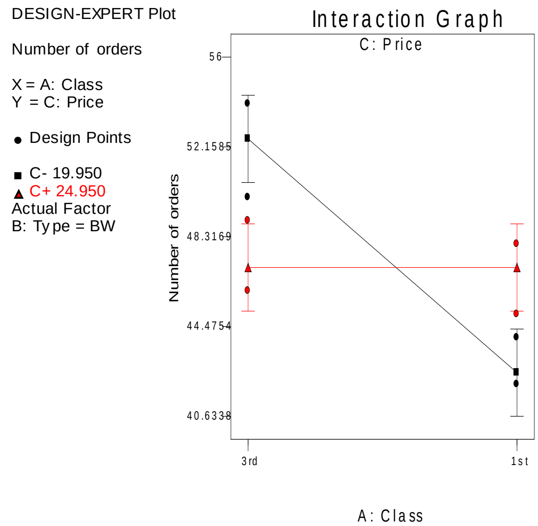


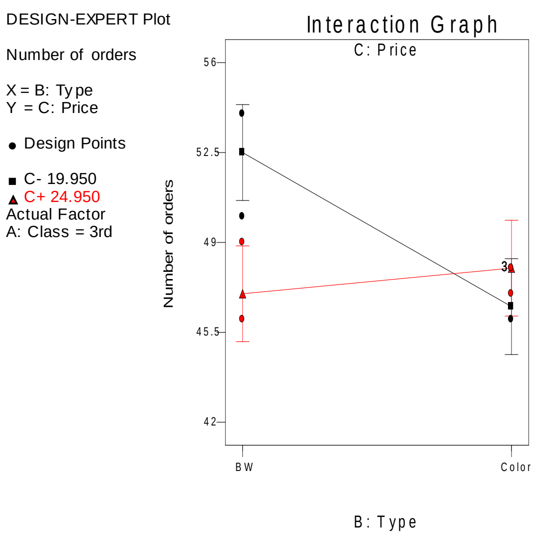


(c) What would you recommend to the company?

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.







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 Ratein 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.

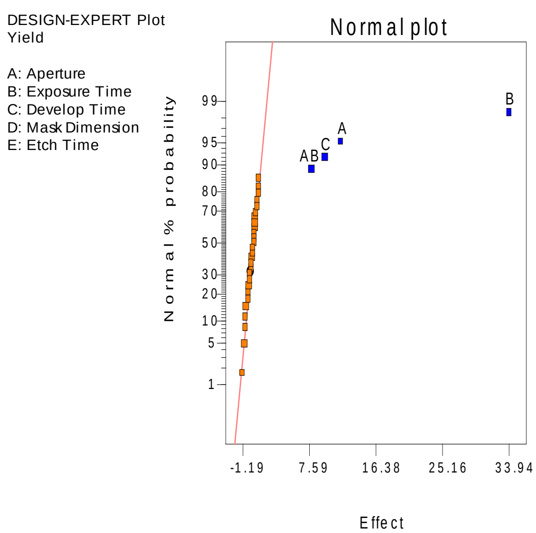
This analysis of variance identifies the same effects as the normal probability plot of effects approach used in Example 6.2. In general, it is not a good idea to arbitrarily pool interactions. Use the normal probability plot of effect estimates as a guide in the choice of which effects to tentatively 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	*d* =	8	*e* =	8	*de* =	6
a =	9	*ad* =	10	*ae* =	12	*ade* =	10
b =	34	*bd* =	32	*be* =	35	*bde* =	30
ab =	55	*abd* =	50	*abe* =	52	*abde* =	53
c =	16	*cd* =	18	*ce* =	15	*cde* =	15
ac =	20	*acd* =	21	*ace* =	22	*acde* =	20
bc =	40	*bcd* =	44	*bce* =	45	*bcde* =	41
abc =	60	*abcd* =	61	*abce* =	65	*abcde* =	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	significant
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:

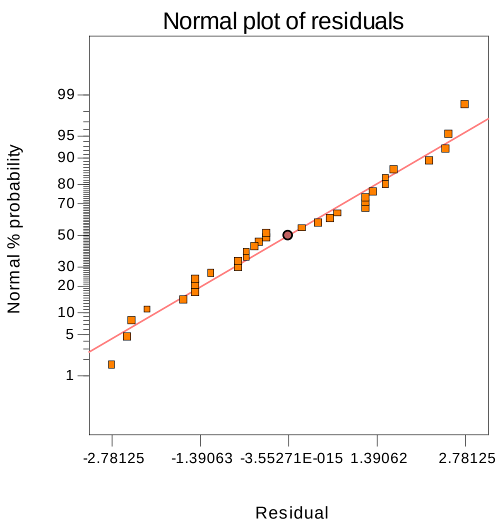
Aperture small

Yield =
+0.40625
+0.65000 * Exposure Time
+0.64583 * Develop Time

Aperture large

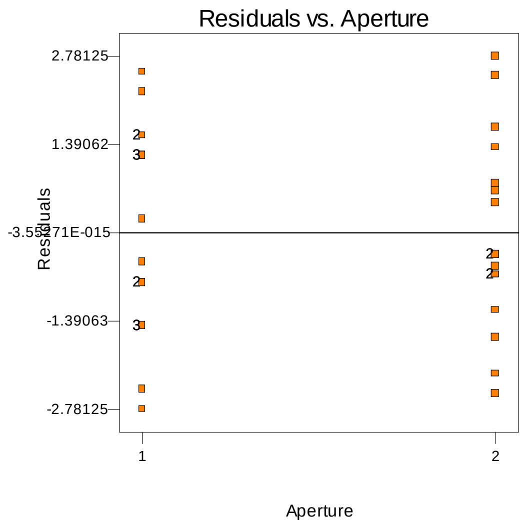
Yield =
+12.21875
+1.04688 * Exposure Time
+0.64583 * Develop Time

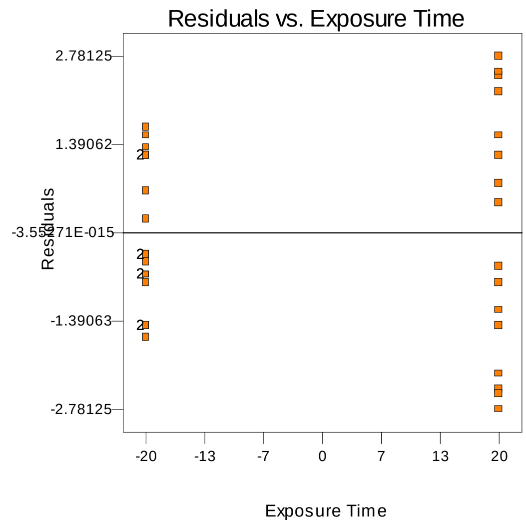
(d) Plot the residuals on normal probability paper. Is the plot satisfactory?

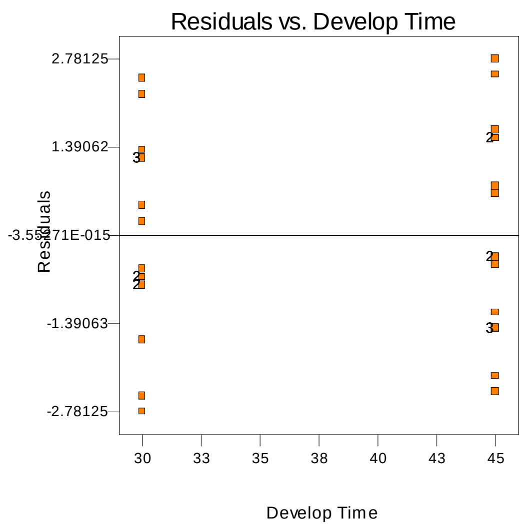


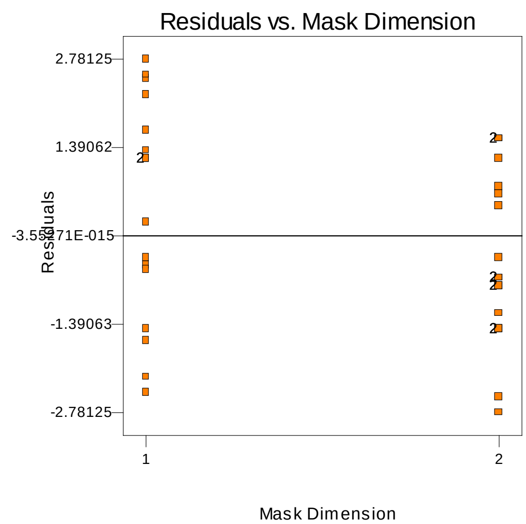
There is nothing unusual about this plot.

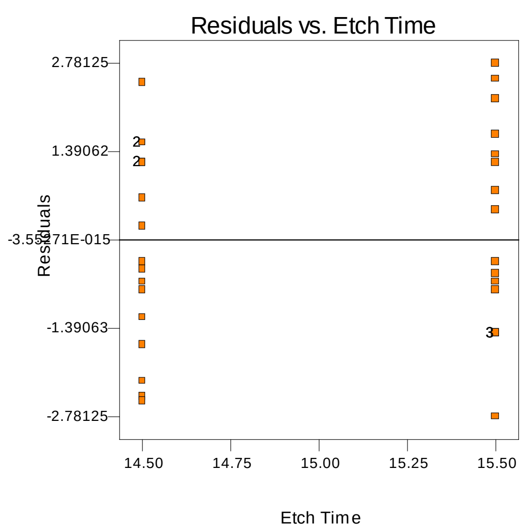
(e) Plot the residuals versus the predicted yields and versus each of the five factors. Comment on the plots.


Residuals vs. Aperture
Residuals
2.78125
1.39062
-3.55271E-015
-1.39063
-2.78125
1
2
Aperture


Residuals vs. Exposure Time
Residuals
2.78125
1.39062
-3.55271E-015
-1.39063
-2.78125
-20
-13
-7
0
7
13
20
Exposure Time

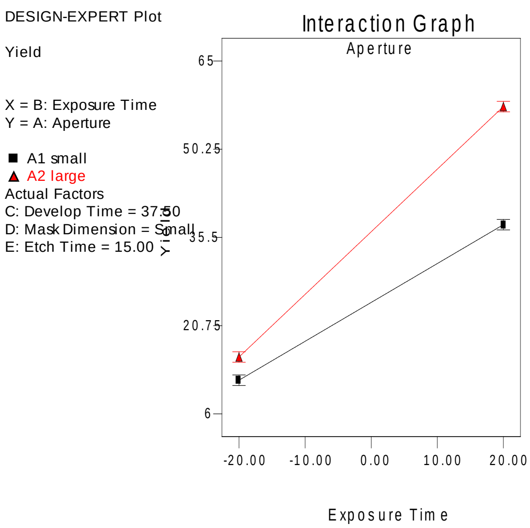

Residuals vs. Develop Time
Residuals
2.78125
1.39062
-3.55271E-015
-1.39063
-2.78125
30
33
35
38
40
43
45
Develop Time


Residuals vs. Mask Dimension
Residuals
2.78125
1.39062
-3.55271E-015
-1.39063
-2.78125
1
2
Mask Dimension


Residuals vs. Etch Time
Residuals
2.78125
1.39062
-3.55271E-015
-1.39063
-2.78125
14.50
14.75
15.00
15.25
15.50
Etch Time

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.

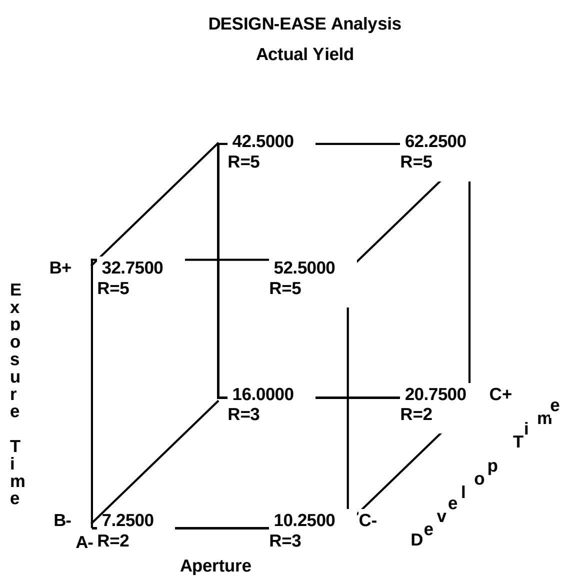


Factor *A* does not have as large an effect when *B* is at its low level as it does when *B* is at its high level.

(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.

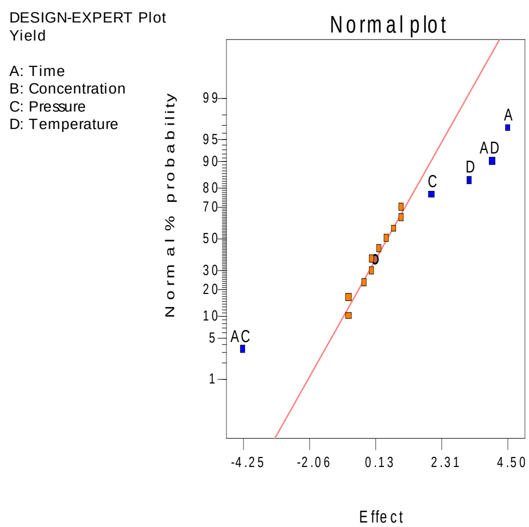
Discuss what your next step would be.

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 Low (-)	Levels 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

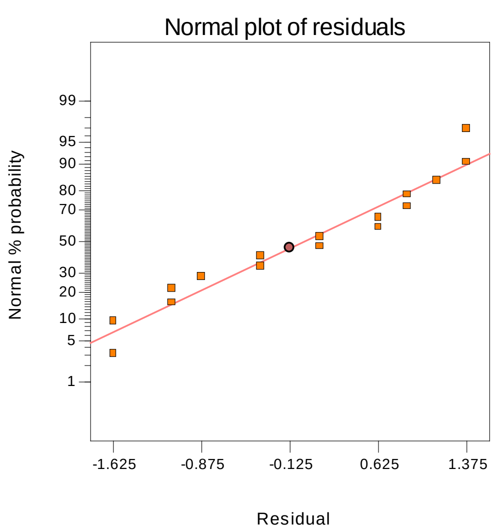
Final Equation in Terms of Coded Factors:

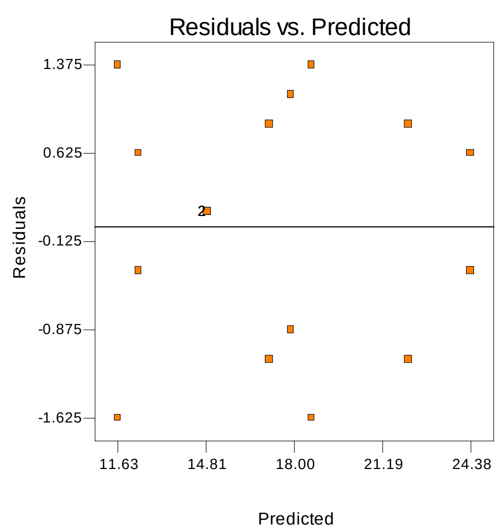
Yield	=
+17.38	
+2.25	*A
+1.00	*C
+1.63	*D
-2.13	*A*C
+2.00	*A*D

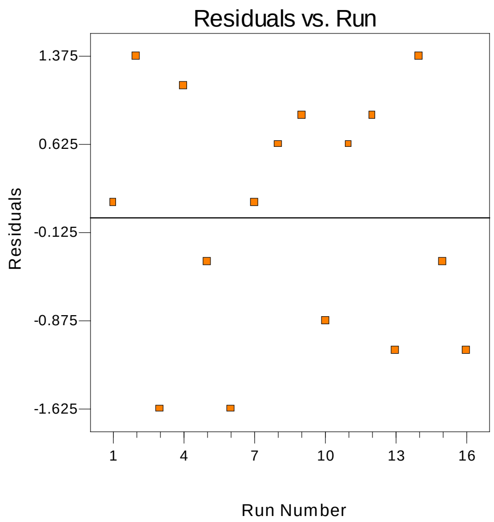
Final Equation in Terms of Actual Factors:

Yield	=
+209.12500	
-83.50000	* Time
+2.43750	* Pressure
-1.63000	* Temperature
-0.85000	* Time * Pressure
+0.64000	* Time * Temperature

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

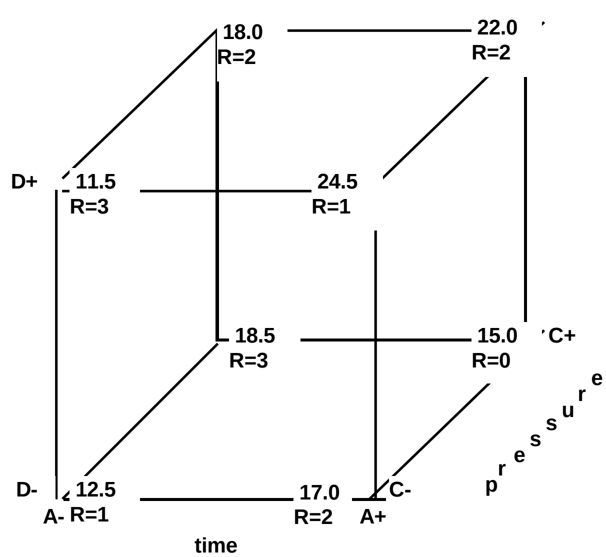






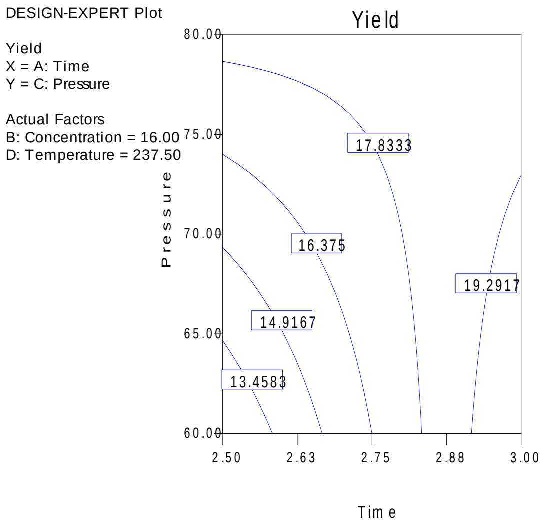
There is nothing unusual about the residual plots.

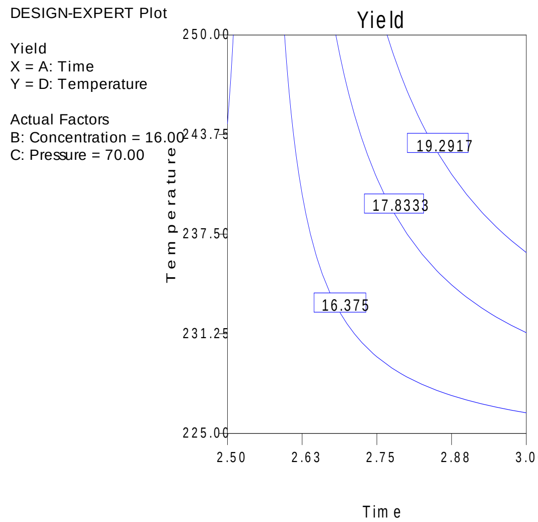
(e) Can this design be collapsed into a 2^3 design with two replicates? If so, sketch the design with the average and range of yield shown at each point in the cube. Interpret the results.



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					Test	Panel	Results				
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.

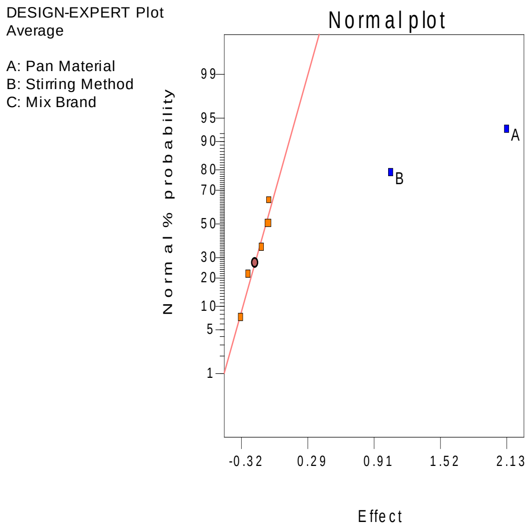
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A are significant model terms.

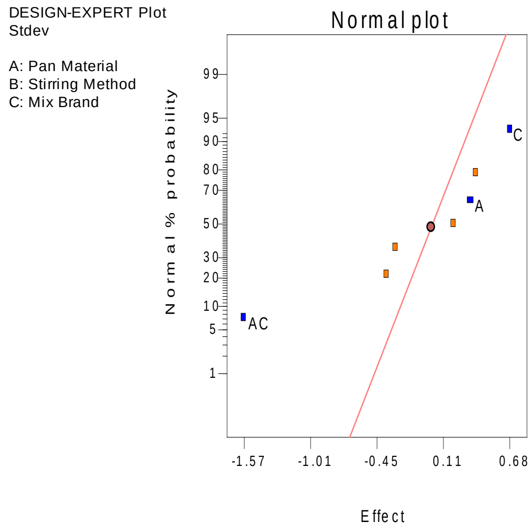
In this analysis, *A*, the pan material and *B*, the stirring method, appear to be significant. There are 56 degrees of freedom for the error, yet only eight batches of brownies were cooked, one for each recipe.

(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 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.

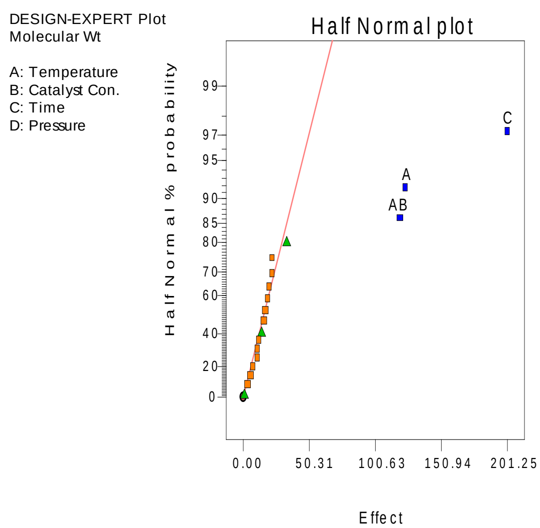
Variables *A* and *B* affect the mean rank of the brownies. Note that the *AC* interaction affects the standard deviation of the ranks. This is an indication that both factors *A* and *C* have some effect on the variability in the ranks. It may also indicate that there is some inconsistency in the taste 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 Order	*A*	*B*	*C*	*D*	Molecular Weight	Viscosity		Factor Low (-)	Levels High (+)
1	18	-	-	-	-	2400	1400	*A* (°C)	100	120
2	9	+	-	-	-	2410	1500	*B* (%)	4	8
3	13	-	+	-	-	2315	1520	*C* (min)	20	30
4	8	+	+	-	-	2510	1630	*D* (psi)	60	75
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.

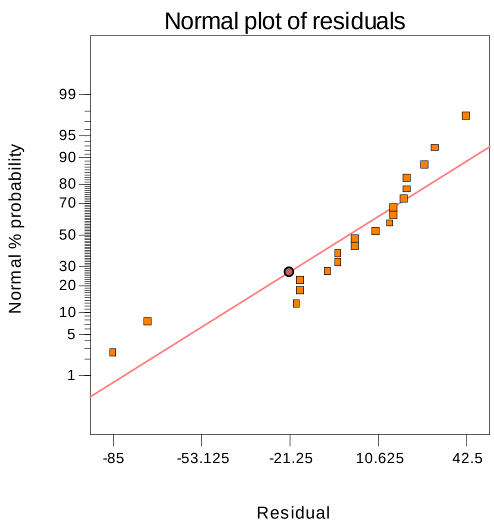
(c) Write down a regression model to predict molecular weight as a function of the important variables.

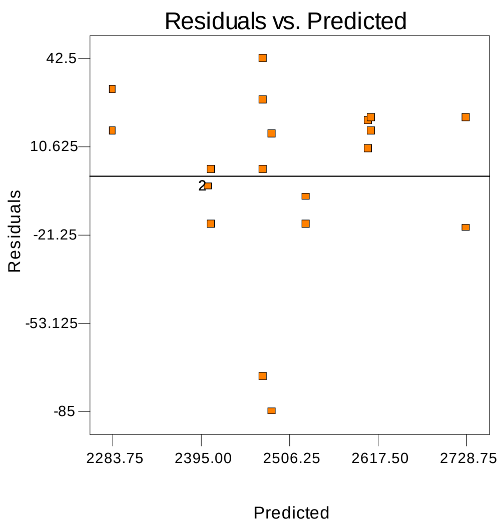
Design Expert Output

Final Equation in Terms of Coded Factors:

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

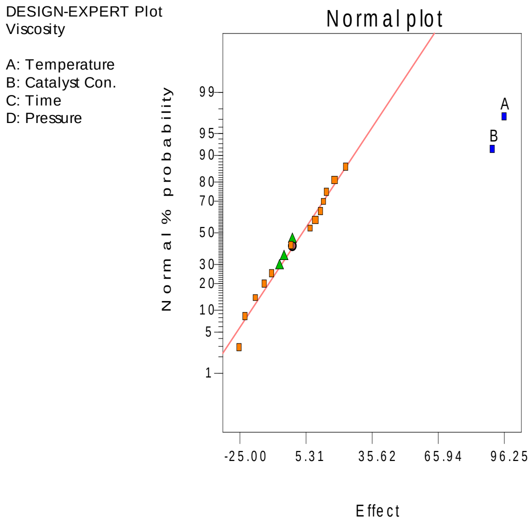
(d) Analyze the residuals and comment on model adequacy.





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]

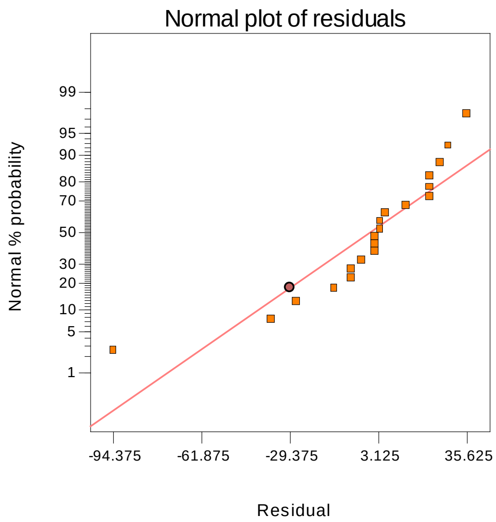
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	70362.50	2	35181.25	35.97	< 0.0001	significant
A	*37056.25*	*1*	*37056.25*	*37.88*	*< 0.0001*	
B	*33306.25*	*1*	*33306.25*	*34.05*	*< 0.0001*	
Curvature	61.25	1	61.25	0.063	0.8056	not significant
Residual	15650.00	16	978.13			
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				

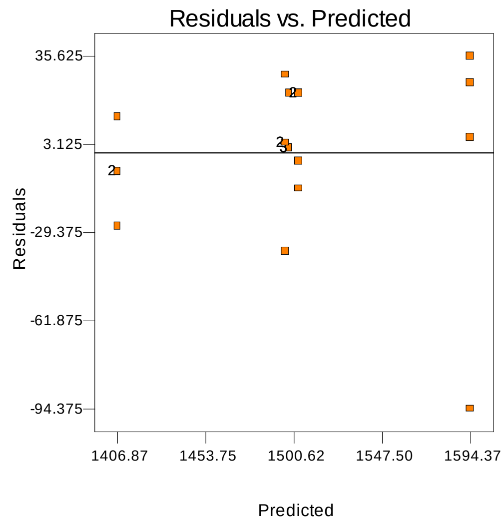
The Model F-value of 35.97 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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:

Viscosity =
+1500.62
+48.13 * A
+45.63 * B

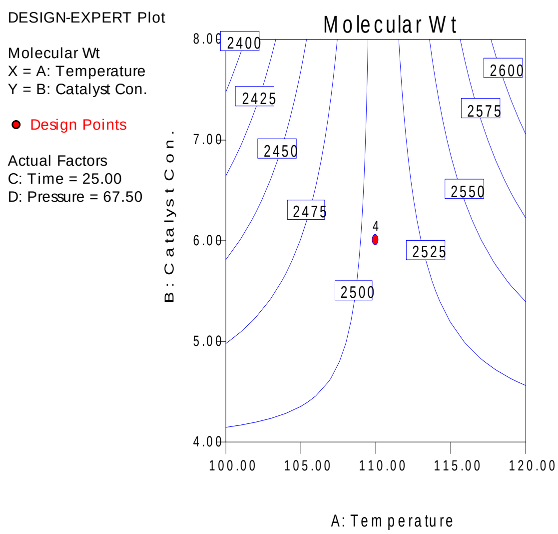




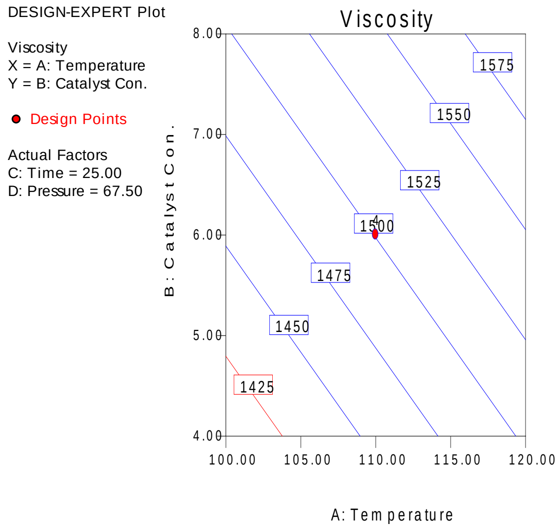
There is one large residual that should be investigated.

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.

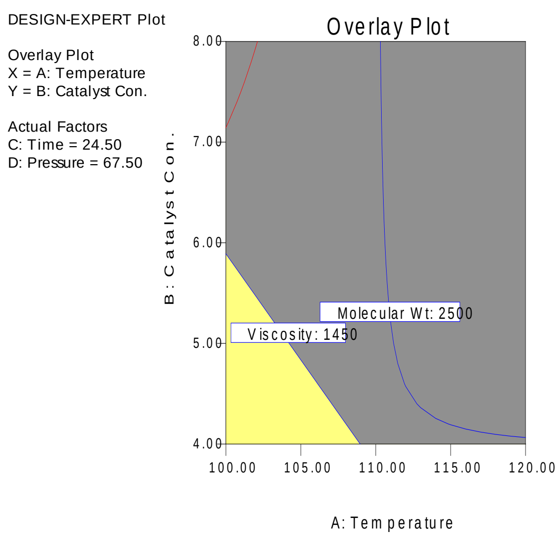


(b) Construct a response surface contour plot for viscosity. In what direction would you adjust the process variables to decrease viscosity?



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.

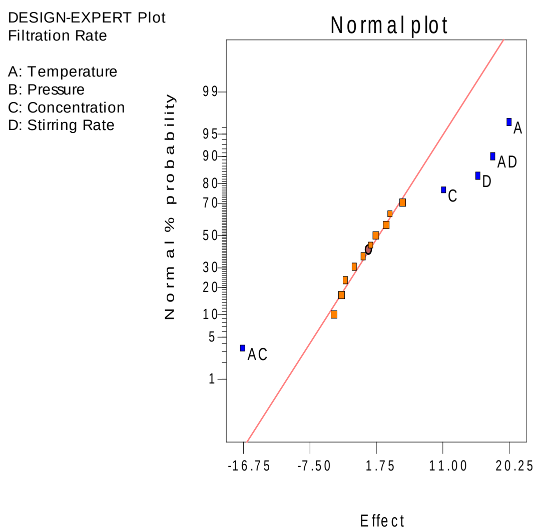
Treatment Combination	Response	Response * *ABCD*	*ABCD*	*A*	*B*	*C*	*D*
(1)	45	45	1	-1	-1	-1	-1
a	71	-71	-1	1	-1	-1	-1
b	48	-48	-1	-1	1	-1	-1
ab	missing	missing * 1	1	1	1	-1	-1
c	68	-68	-1	-1	-1	1	-1
ac	60	60	1	1	-1	1	-1
bc	80	80	1	-1	1	1	-1
abc	65	-65	-1	1	1	1	-1
d	43	-43	-1	-1	-1	-1	1
ad	100	100	1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

Contrast (ABCD) = Missing – 54 = 0
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		

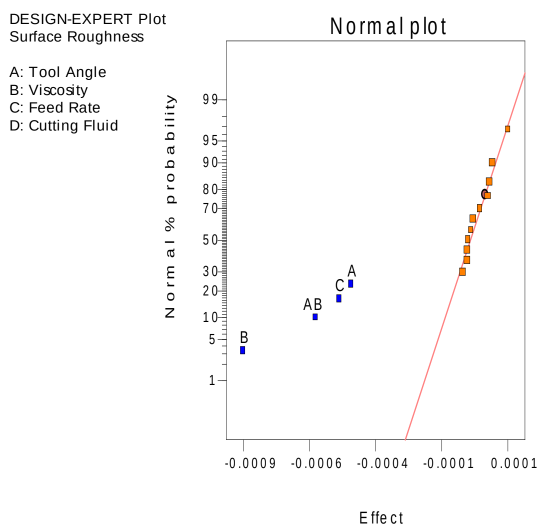


6.35. An engineer has performed an experiment to study the effect of four factors on the surface roughness of a machined part. The factors (and their levels) are *A* = tool angle (12 degrees, 15 degrees), *B* = cutting fluid viscosity (300, 400), *C* = feed rate (10 in/min, 15 in/min), and *D* = cutting 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

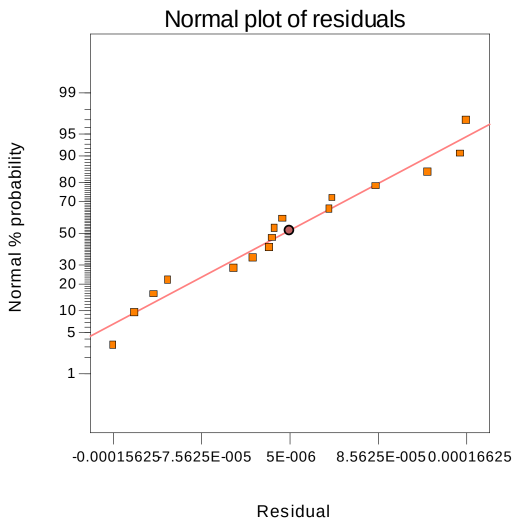
Response: Surface Roughness
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

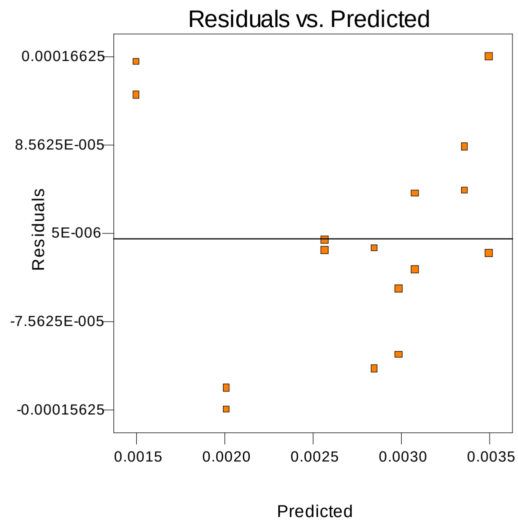
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	6.406E-006	4	1.601E-006	114.97	< 0.0001	significant
A	*8.556E-007*	*1*	*8.556E-007*	*61.43*	*< 0.0001*	
B	*3.080E-006*	*1*	*3.080E-006*	*221.11*	*< 0.0001*	
C	*1.030E-006*	*1*	*1.030E-006*	*73.96*	*< 0.0001*	
AB	*1.440E-006*	*1*	*1.440E-006*	*103.38*	*< 0.0001*	
Residual	1.532E-007	11	1.393E-008			
Cor Total	6.559E-006	15				

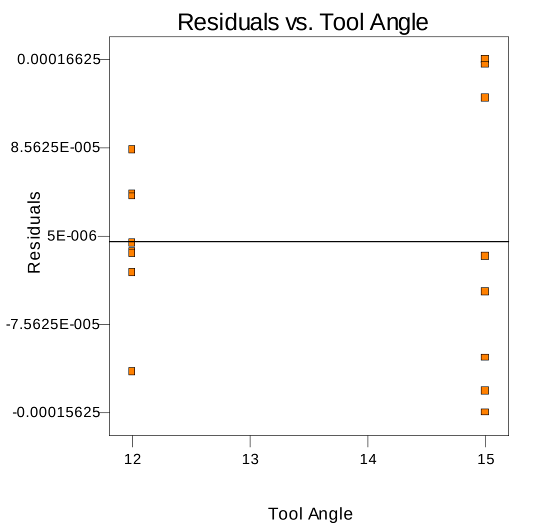
The Model F-value of 114.97 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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

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



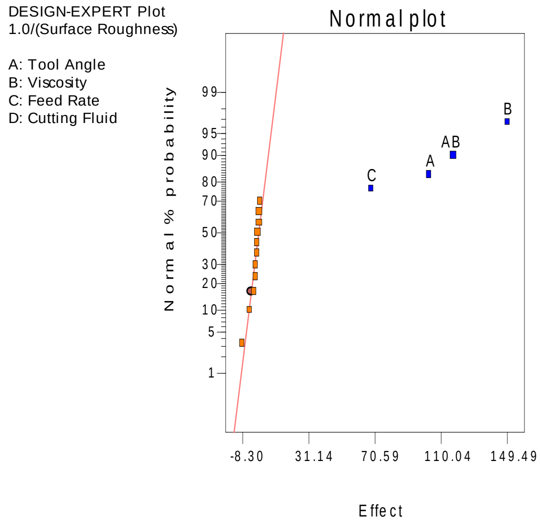




The plot of residuals versus predicted shows a slight "u-shaped" appearance in the residuals, and the plot of residuals versus tool angle shows an outward-opening funnel.

(c) Repeat the analysis from parts (a) and (b) using $1/y$ as the response variable. Is there and 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

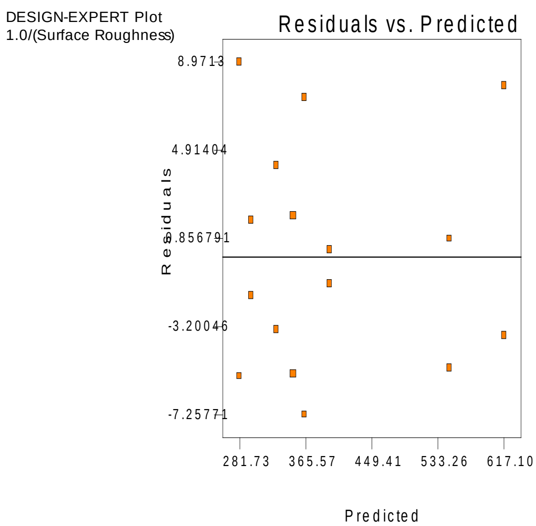
ANOVA for Selected Factorial Model

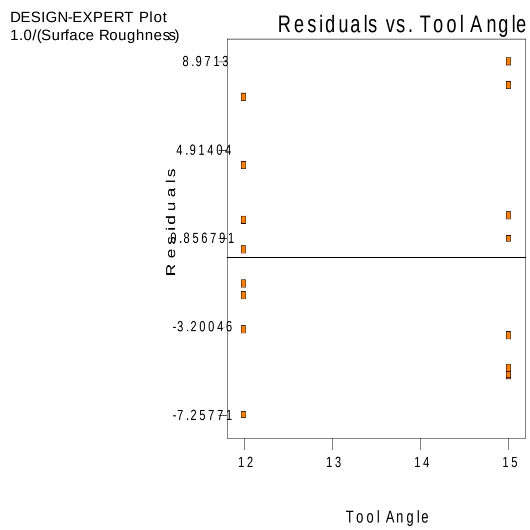
Analysis of variance table [Partial sum of squares]

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

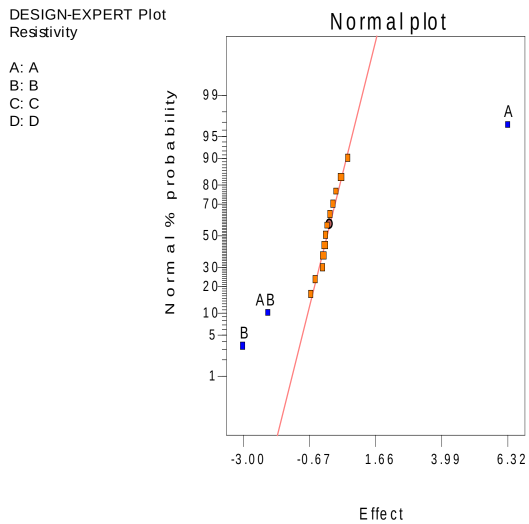
Final Equation in Terms of Coded Factors:

```
1.0/(Surface Roughness)      =
+397.81
+51.61      * A
+74.74      * B
+34.24      * C
+58.70      * A * B
```

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

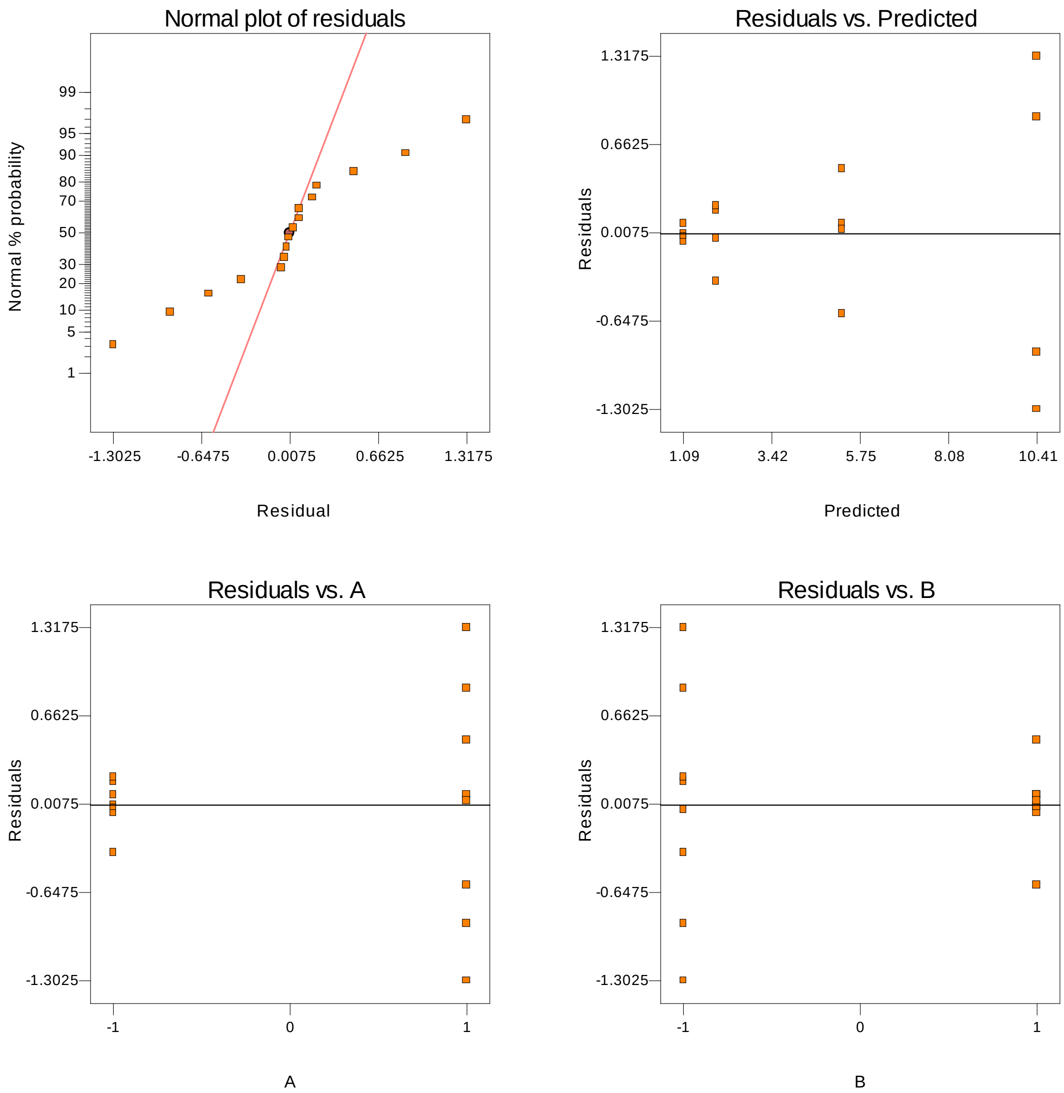
ANOVA for Selected Factorial Model

Analysis of variance table [Partial sum of squares]

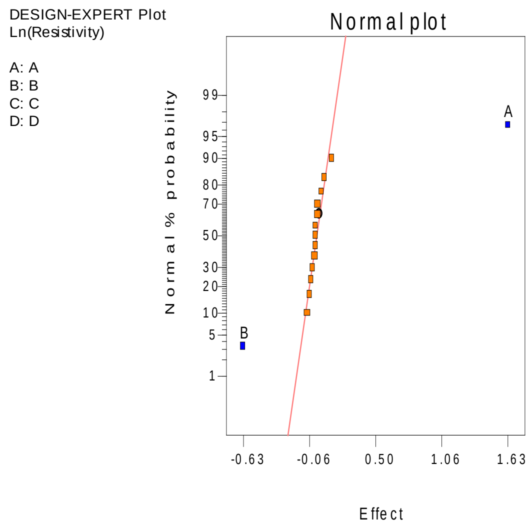
Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	214.22	3	71.41	148.81	< 0.0001	significant
A	*159.83*	*1*	*159.83*	*333.09*	*< 0.0001*	
B	*36.09*	*1*	*36.09*	*75.21*	*< 0.0001*	
AB	*18.30*	*1*	*18.30*	*38.13*	*< 0.0001*	
Residual	5.76	12	0.48			
Cor Total	219.98	15				

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.


Normal plot of residuals
Normal % probability
99
95
90
80
70
50
30
20
10
5
1
-1.3025
-0.6475
0.0075
0.6625
1.3175
Residual
Residuals vs. Predicted
Residuals
1.3175
0.6625
0.0075
-0.6475
-1.3025
1.09
3.42
5.75
8.08
10.41
Predicted
Residuals vs. A
Residuals
1.3175
0.6625
0.0075
-0.6475
-1.3025
-1
0
1
A
Residuals vs. B
Residuals
1.3175
0.6625
0.0075
-0.6475
-1.3025
-1
0
1
B

(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

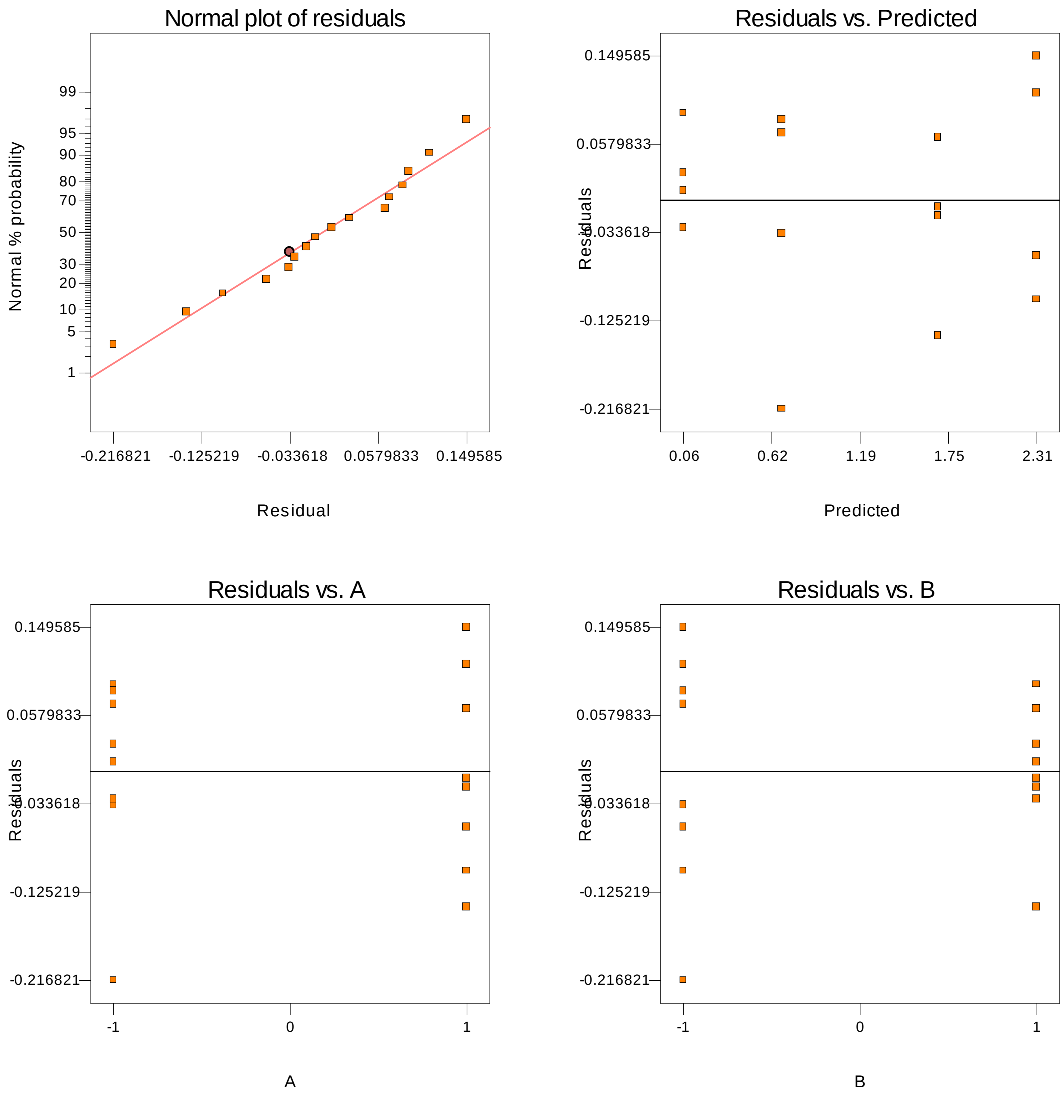
Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F 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.

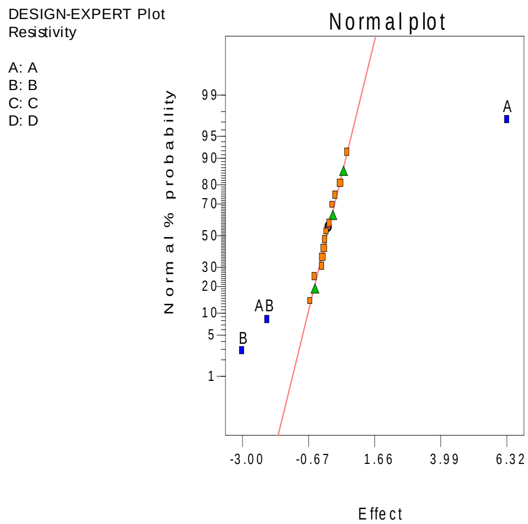
(d) Fit a model in terms of the coded variables that can be used to predict the resistivity.

Design Expert Output

Final Equation in Terms of Coded Factors:

Ln(Resistivity)		=
+1.19		
+0.81	* A	
-0.31	* B	

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

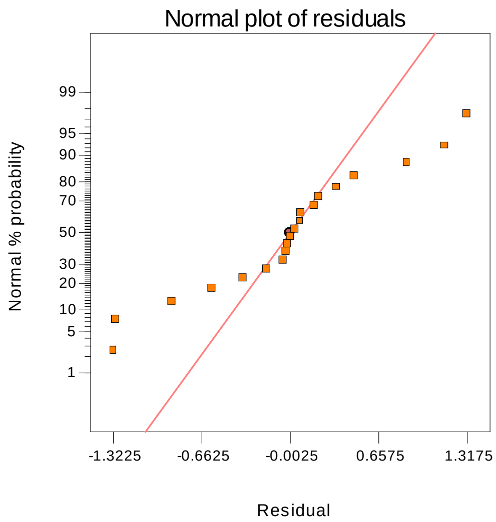
ANOVA for Selected Factorial Model

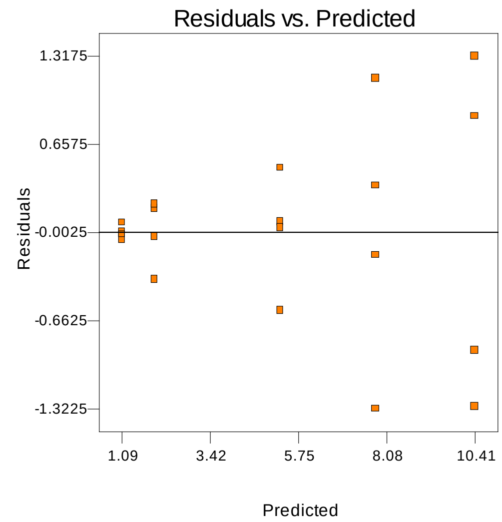
Analysis of variance table [Partial sum of squares]

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				

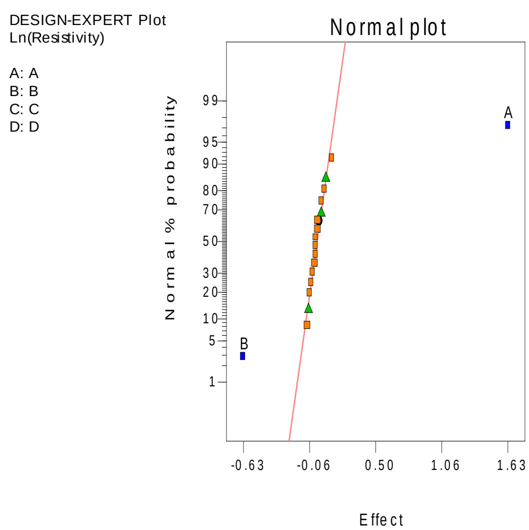
The Model F-value of 119.35 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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 studies 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*	

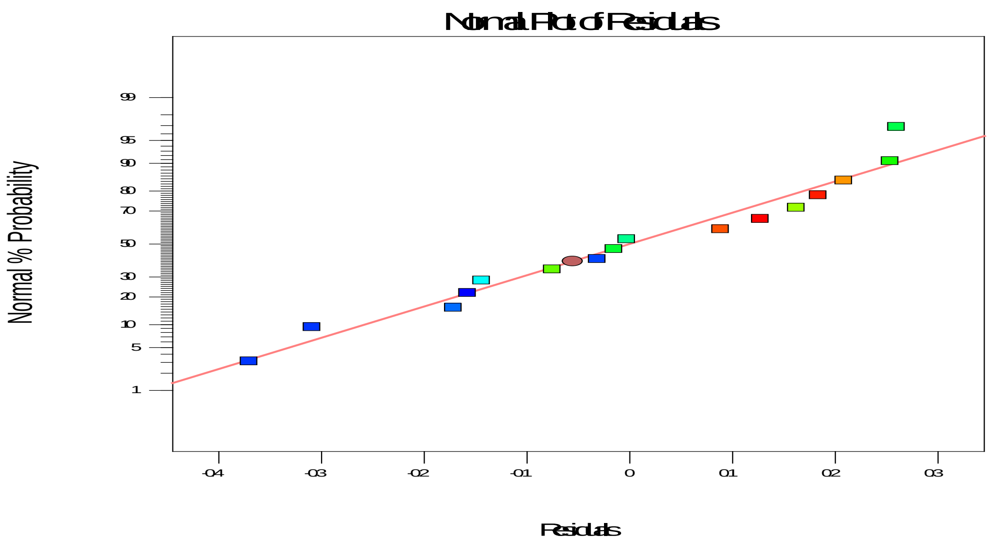
C-Amount of acid	*0.023*	*1*	*0.023*	*0.45*	*0.5181*
D-Temperature	*0.30*	*1*	*0.30*	*5.75*	*0.0354*
Residual	0.57	11	0.052		
Cor Total	1.04	15			

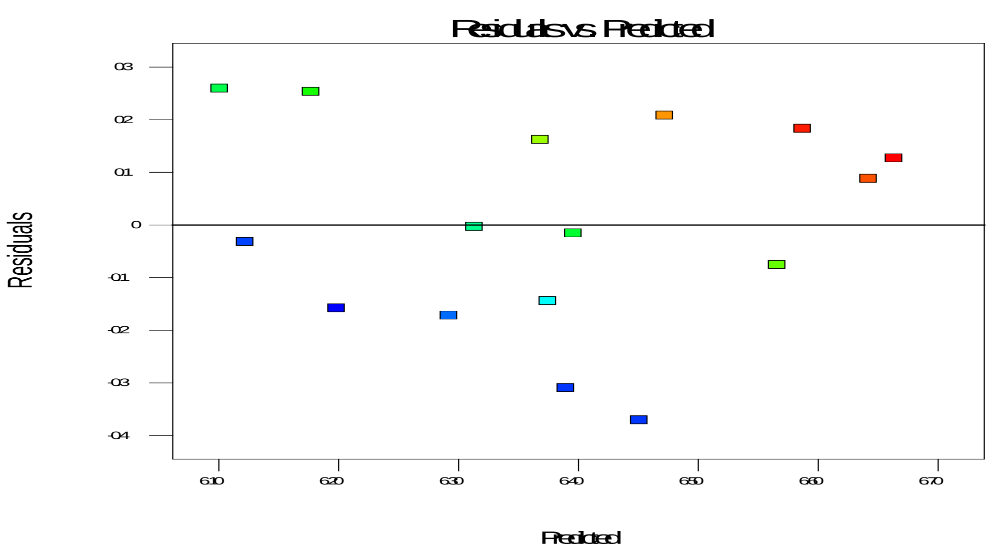
The "Model F-value" of 2.26 implies the model is not significant relative to the noise. There is a 12.87 % chance that a "Model F-value" this large could occur due to noise.

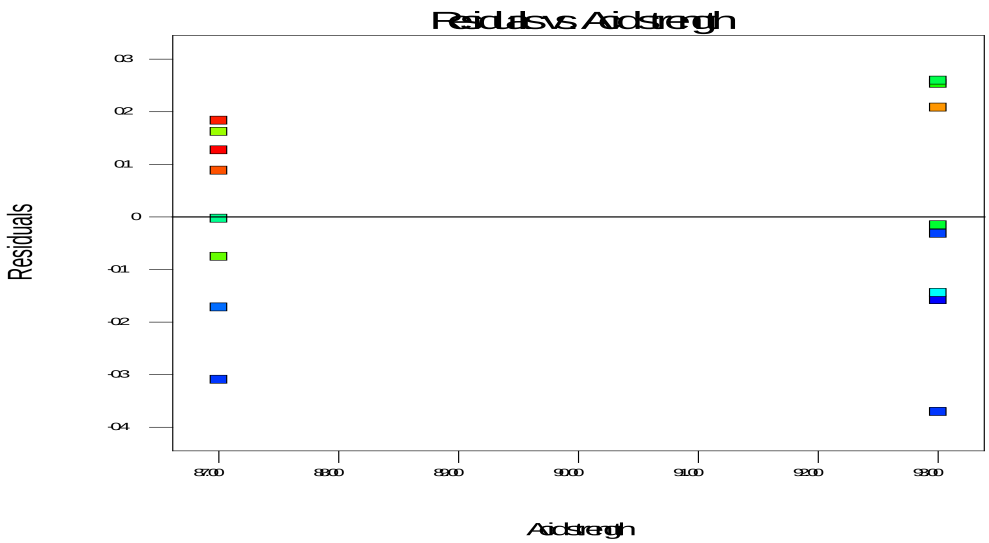
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case D are significant model terms.

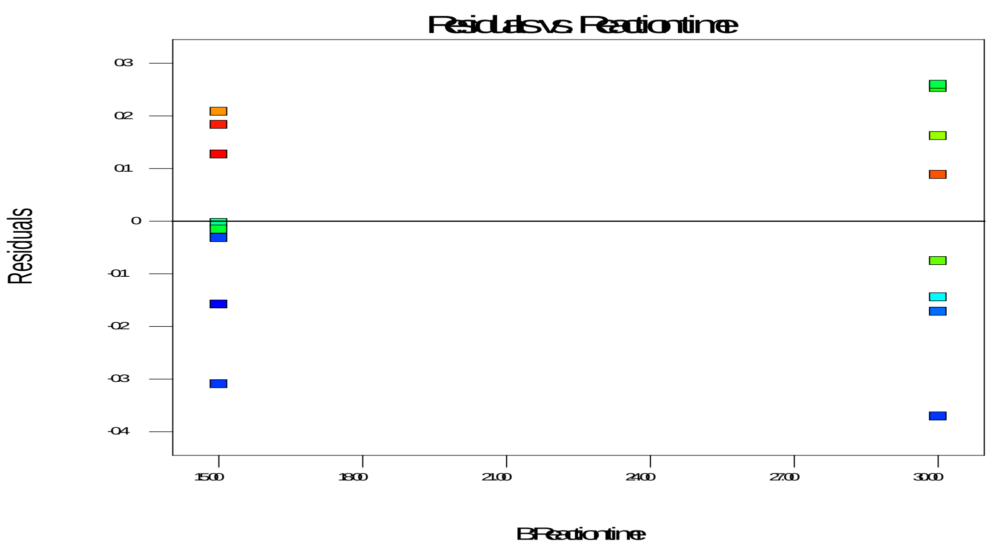
(b) Analyze the residuals. Are there any indications of model inadequacy or violation of the assumptions?

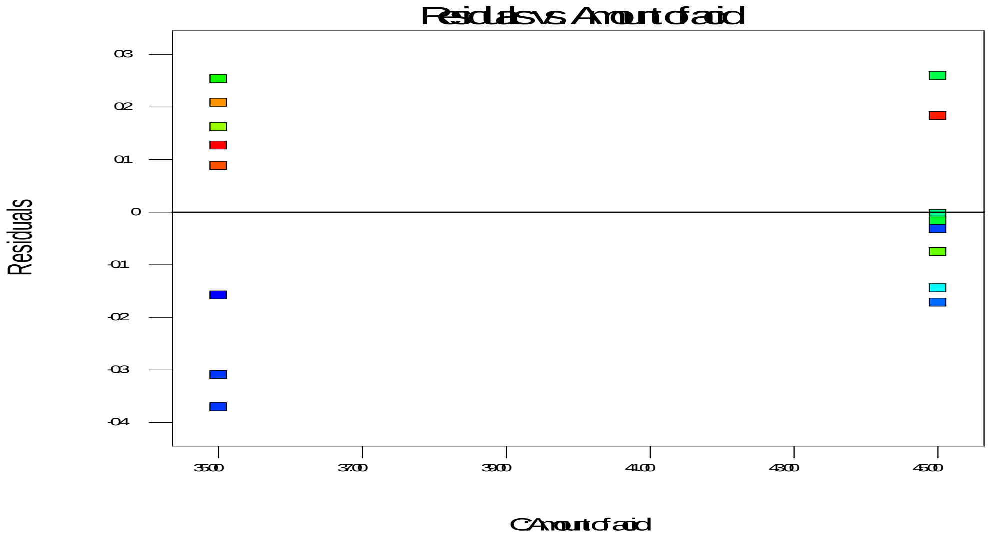
The residual plots appear to be acceptable.

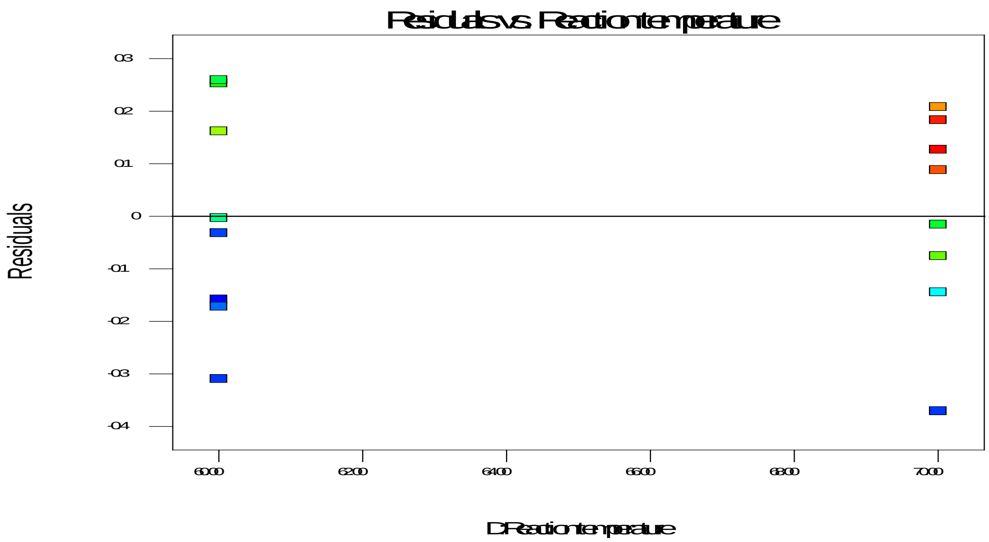





Residuals


Residuals





(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:	
Yield	=
+6.38	
-0.096	* A
-0.011	* B
-0.038	* C
+0.14	* D

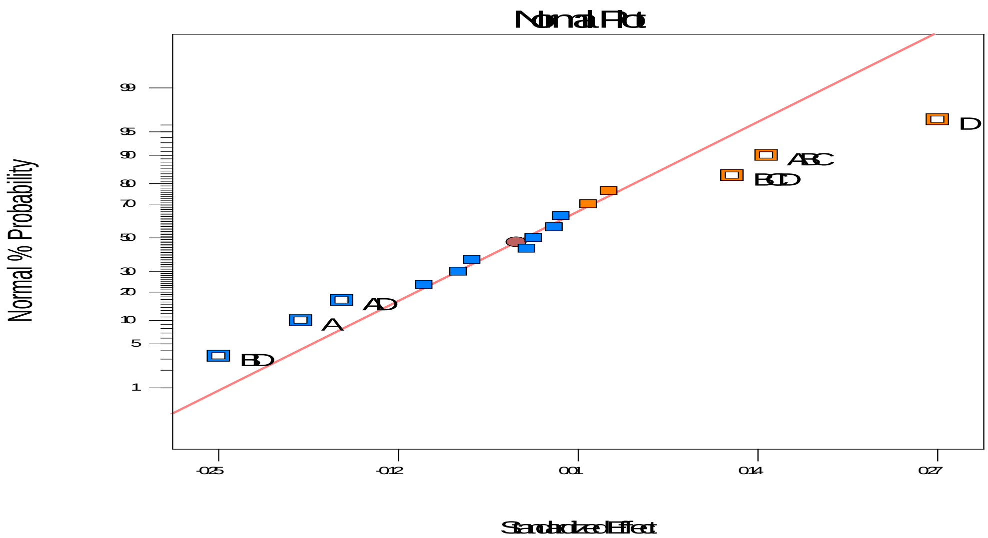
Design Expert Output

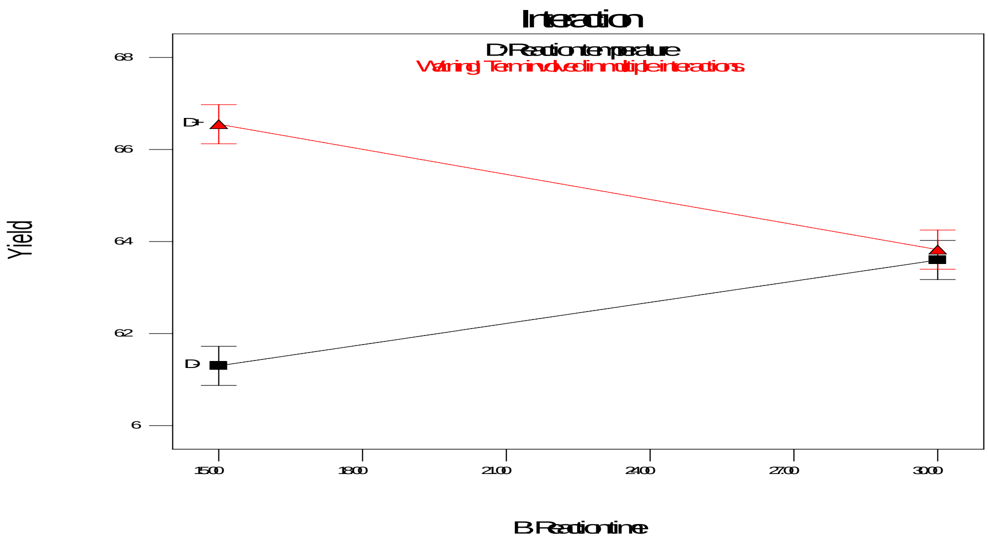
Final Equation in Terms of Coded Factors:	
Yield	=
+7.80813	

-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.



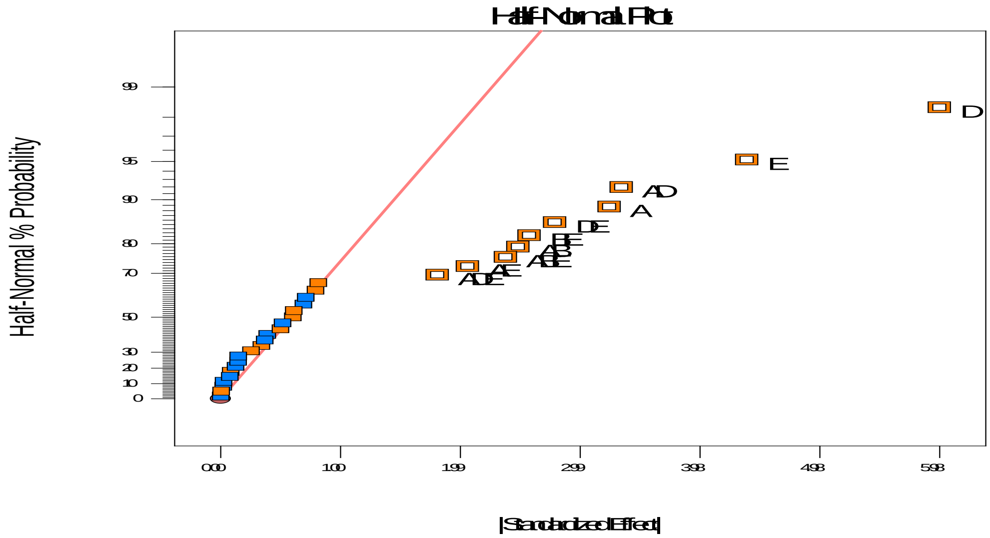


6.39. An article in *Quality and Reliability Engineering International* (2010, Vol. 26, pp. 223-233) presents a 2^5 factorial design. The experiment is shown in the following table:

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]

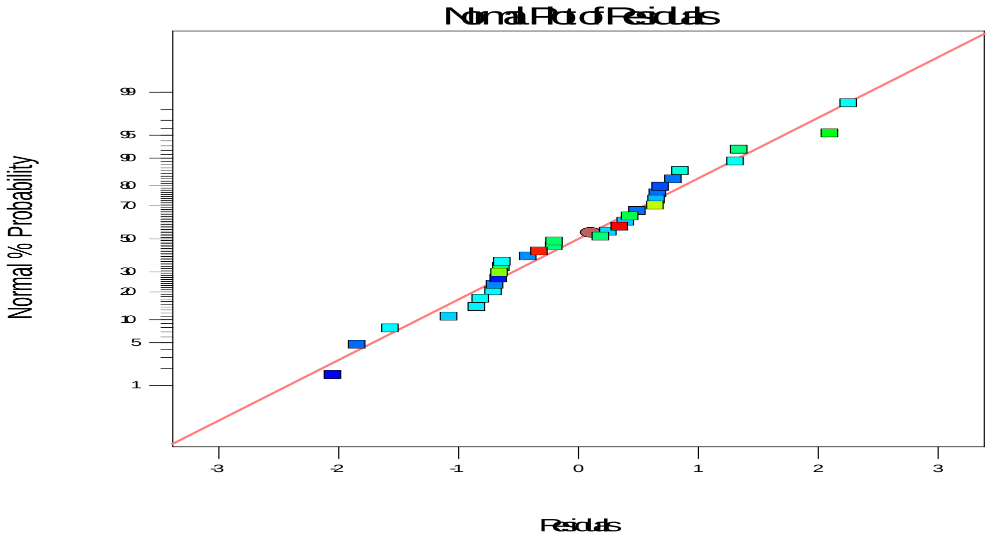
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	879.62	11	79.97	49.15	< 0.0001	significant
A-A	*83.56*	*1*	*83.56*	*51.36*	*< 0.0001*	
B-B	*0.060*	*1*	*0.060*	*0.037*	*0.8492*	
D-D	*285.78*	*1*	*285.78*	*175.66*	*< 0.0001*	
E-E	*153.17*	*1*	*153.17*	*94.15*	*< 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				

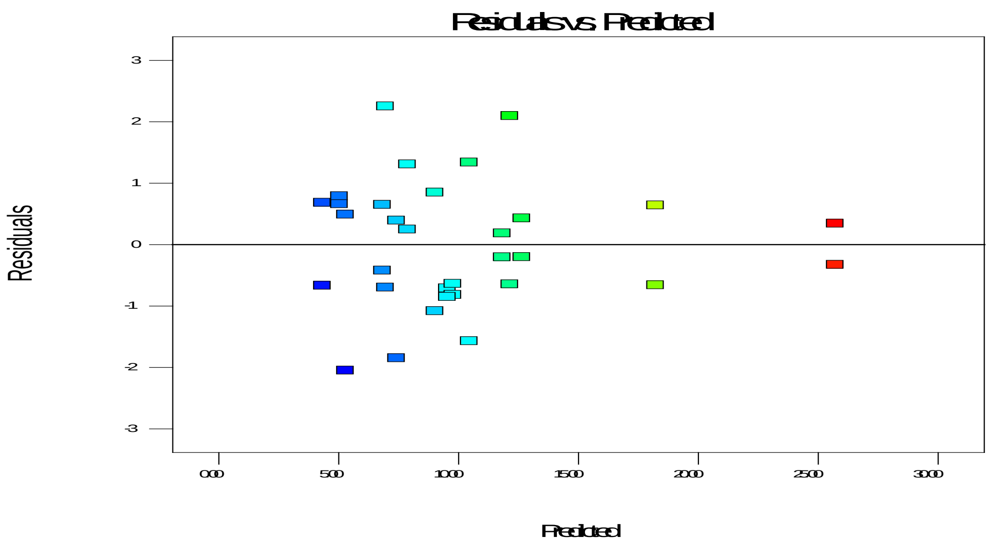
The Model F-value of 49.15 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

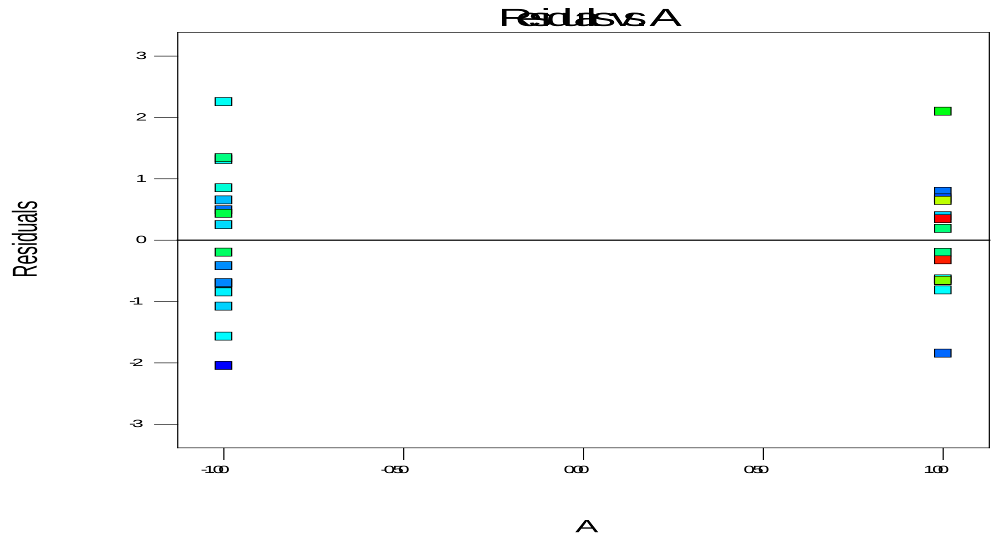
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, D, E, AB, AD, AE, BE, DE, ABE, ADE are significant model terms.

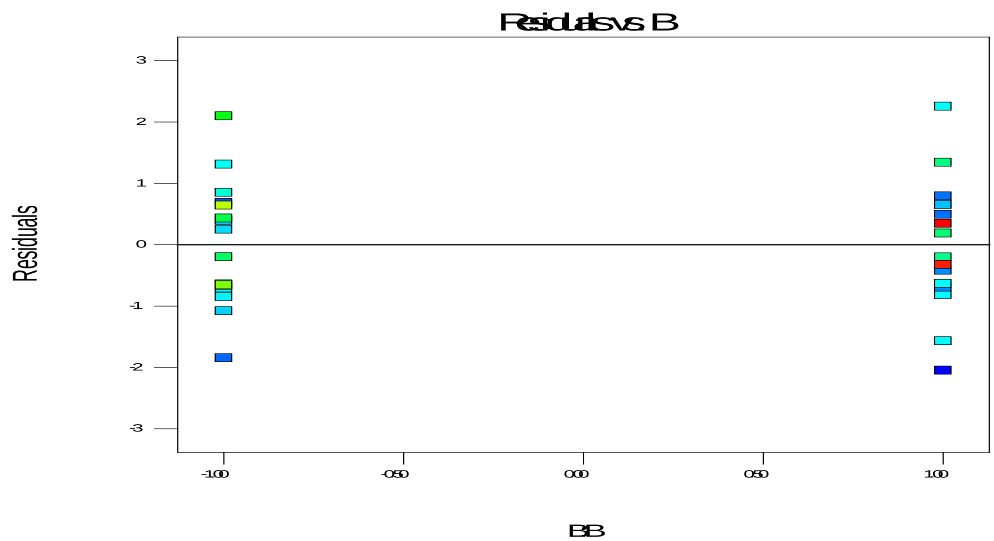
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy or violations of the assumptions?

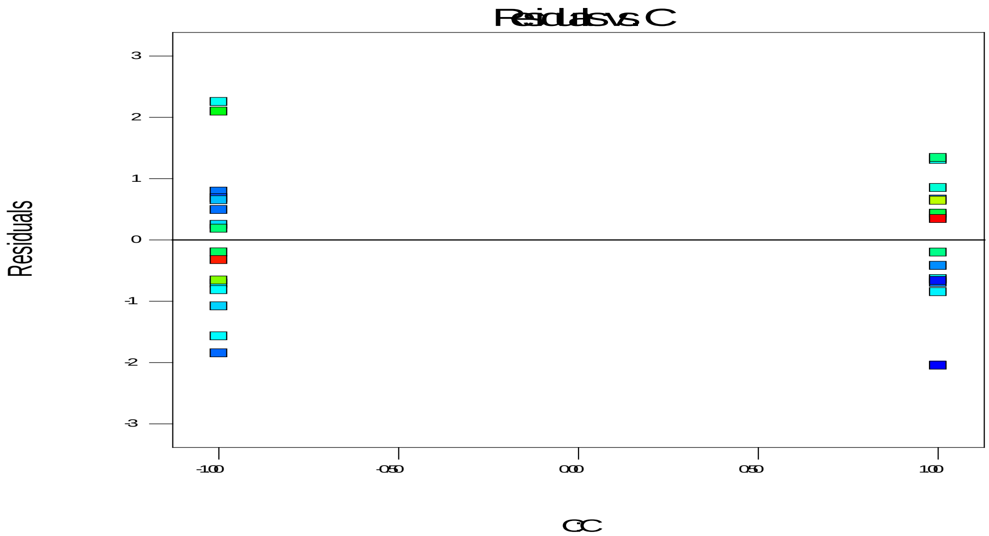
The residual plots below do not identify any concerns with model adequacy or the violations of the assumptions.

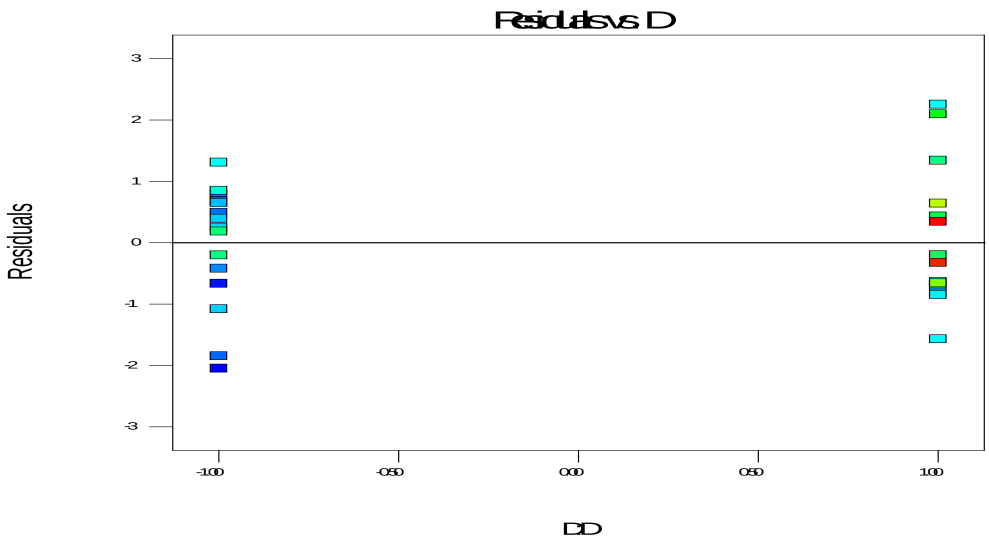

Normal Plot of Residuals
Normal % Probability
Residuals

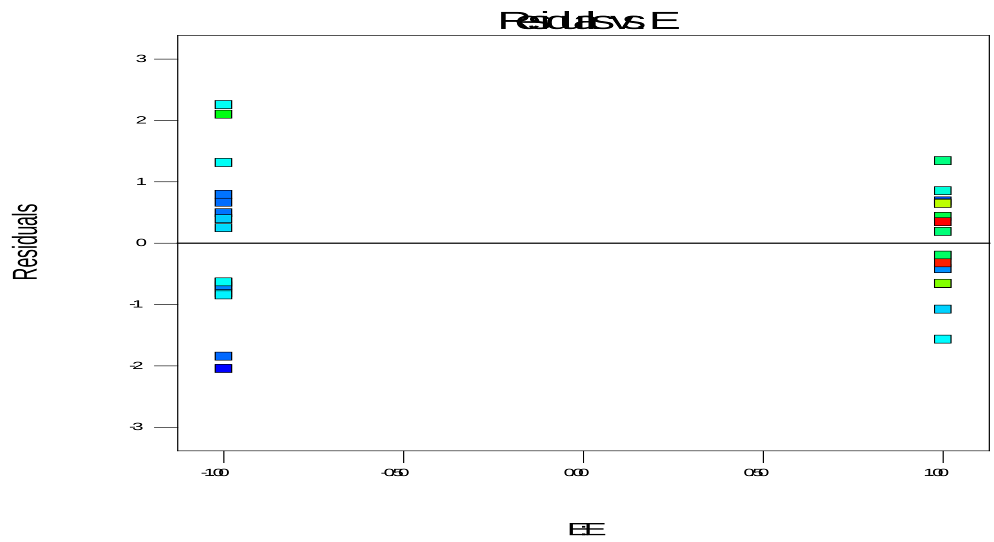

Residuals vs. Predicted
Residuals
Predicted


Residuals vs A
Residuals
3
2
1
0
-1
-2
-3
-1.00
-0.50
0.00
0.50
1.00
A


Residuals vs B
Residuals
3
2
1
0
-1
-2
-3
-1.00
-0.50
0.00
0.50
1.00
B


Residuals vs C
Residuals
C:C


Residuals vs D
Residuals
D:D



(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]

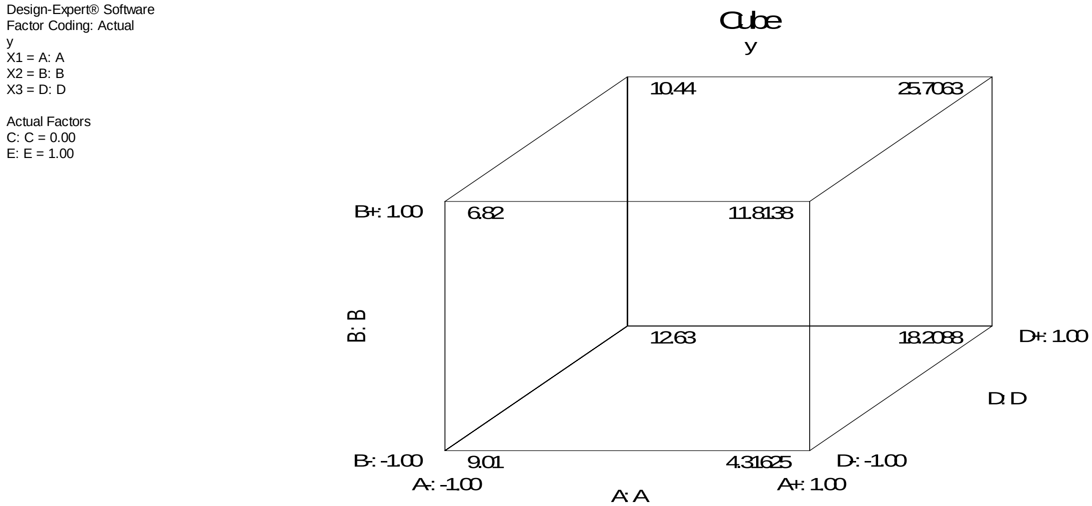
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	888.80	15	59.25	40.58	< 0.0001	significant
A-A	*83.56*	*1*	*83.56*	*57.23*	*< 0.0001*	
B-B	*0.060*	*1*	*0.060*	*0.041*	*0.8414*	
D-D	*285.78*	*1*	*285.78*	*195.74*	*< 0.0001*	
E-E	*153.17*	*1*	*153.17*	*104.91*	*< 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.

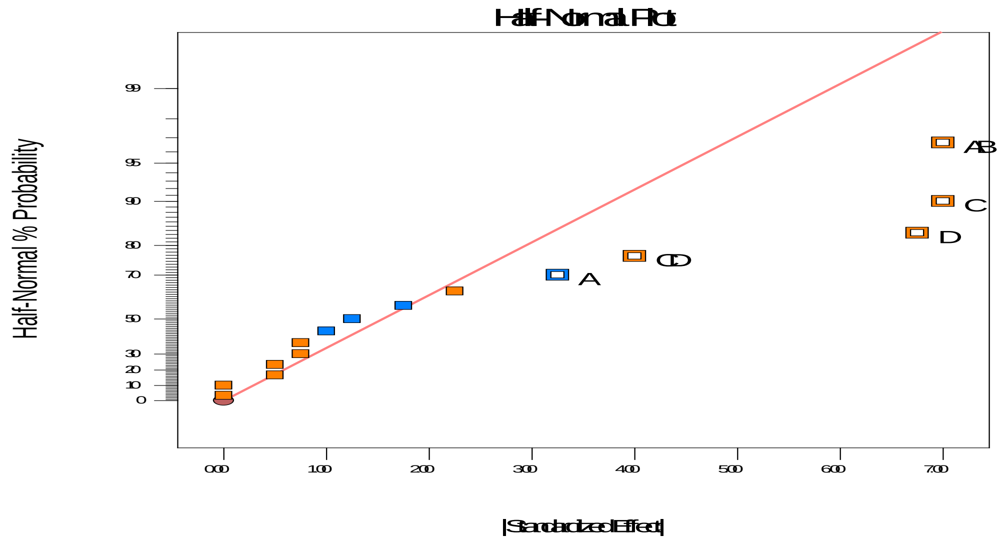


6.40. A paper in the *Journal of Chemical Technology and Biotechnology* ("Response Surface Optimization of the Critical Media Components for the Production of Surfactin," 1997, Vol. 68, pp. 263-270) describes the use of a designed experiment to maximize surfactin production. A portion of the data from 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 ($g\ dm^{-3}$)	$FeSO_4$ ($g\ dm^{-3} \times 10^{-4}$)	$MnSO_4$ ($g\ dm^{-3} \times 10^{-2}$)	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]

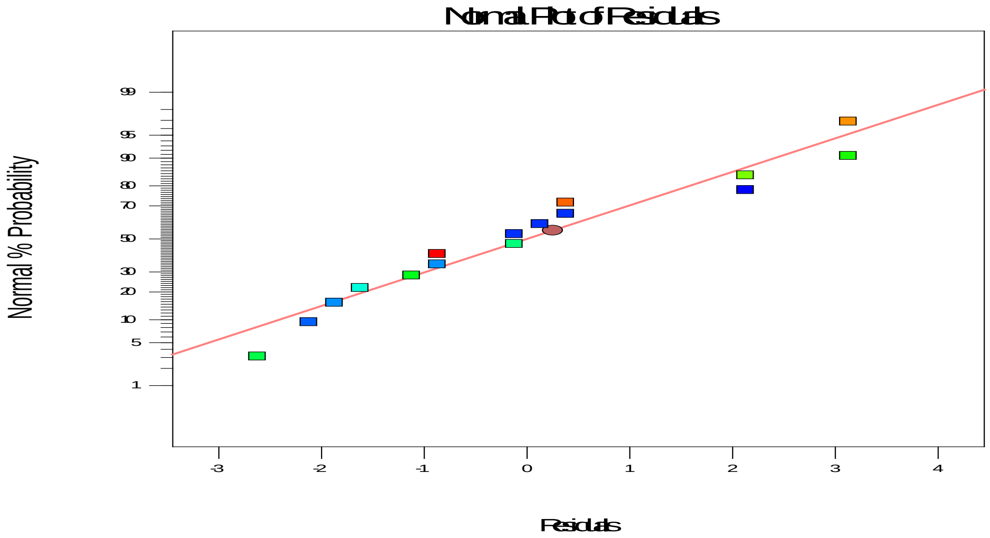
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	680.50	6	113.42	20.73	< 0.0001	significant
A-Glucose	*42.25*	*1*	*42.25*	*7.72*	*0.0214*	
B-NH4NO3	*0.000*	*1*	*0.000*	*0.000*	*1.0000*	
C-FeSO4	*196.00*	*1*	*196.00*	*35.82*	*0.0002*	
D-MnSO4	*182.25*	*1*	*182.25*	*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				

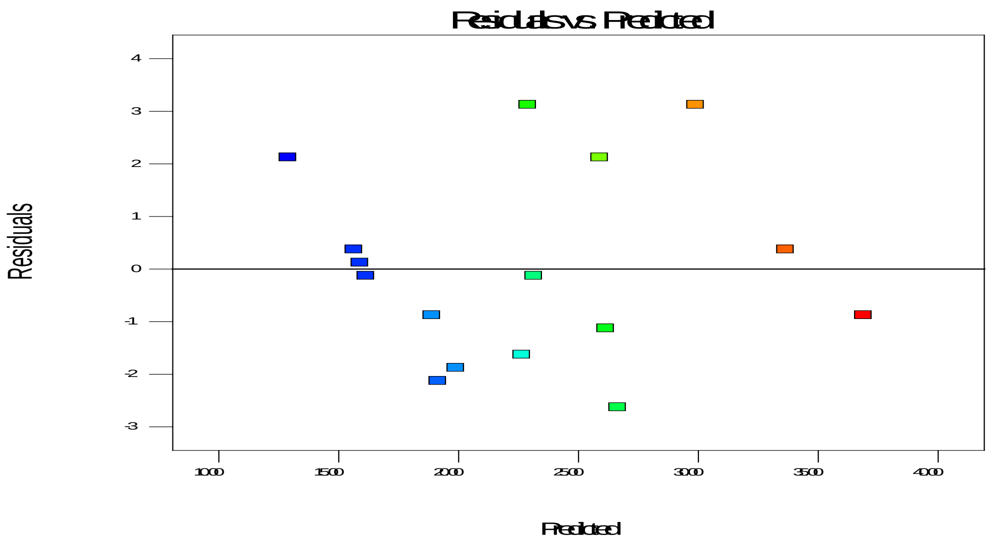
The Model F-value of 20.73 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

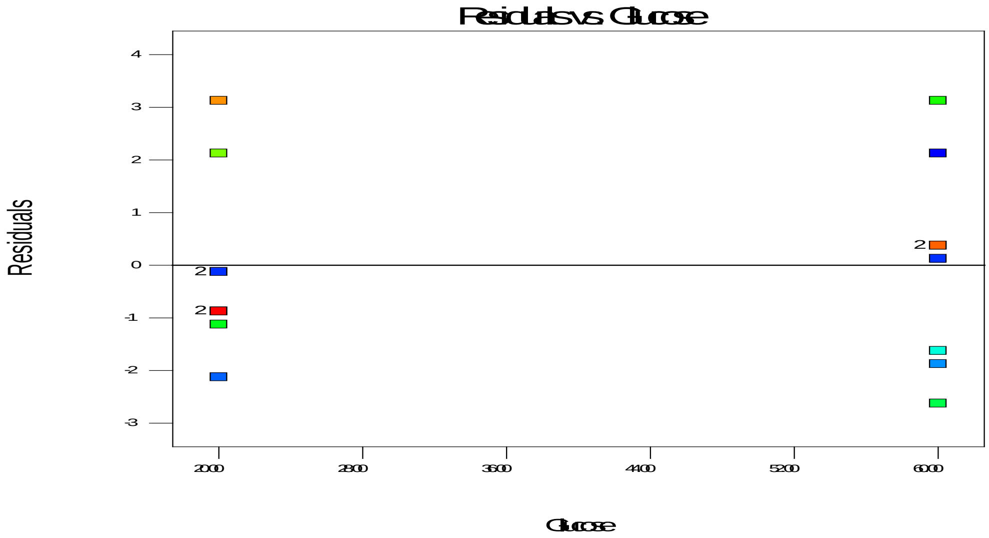
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, C, D, AB, CD are significant model terms.

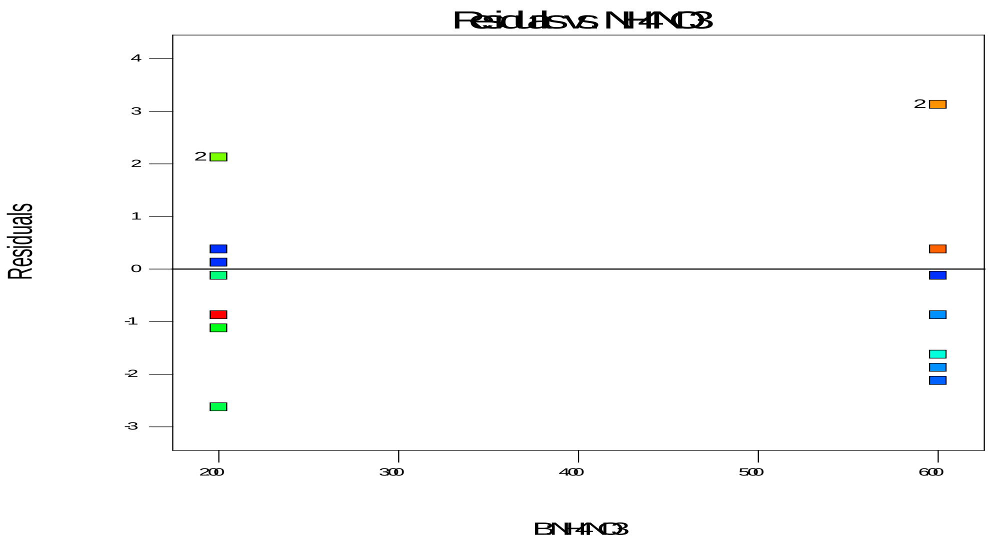
(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy or violations of the assumptions?

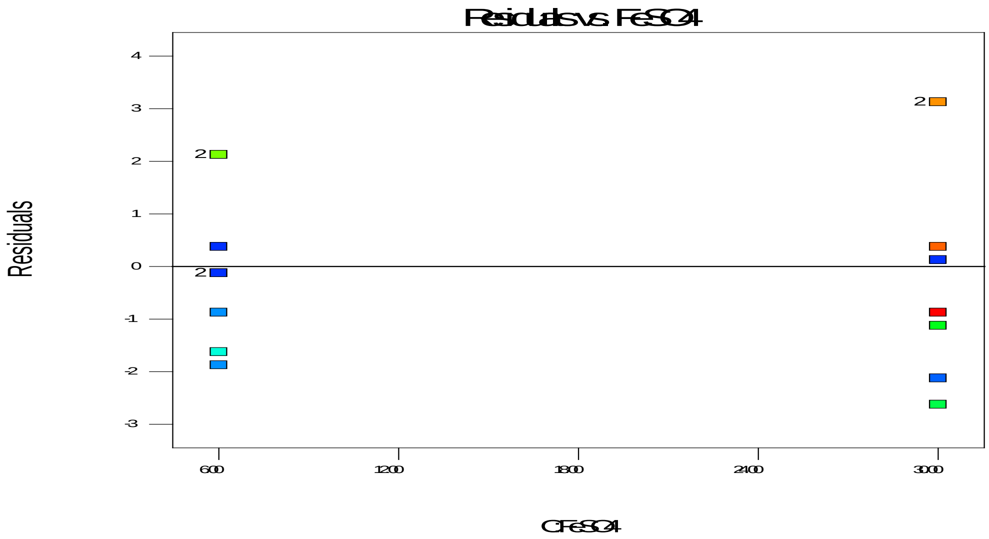
The residual plots below do not identify any concerns with model adequacy or the violations of the assumptions.

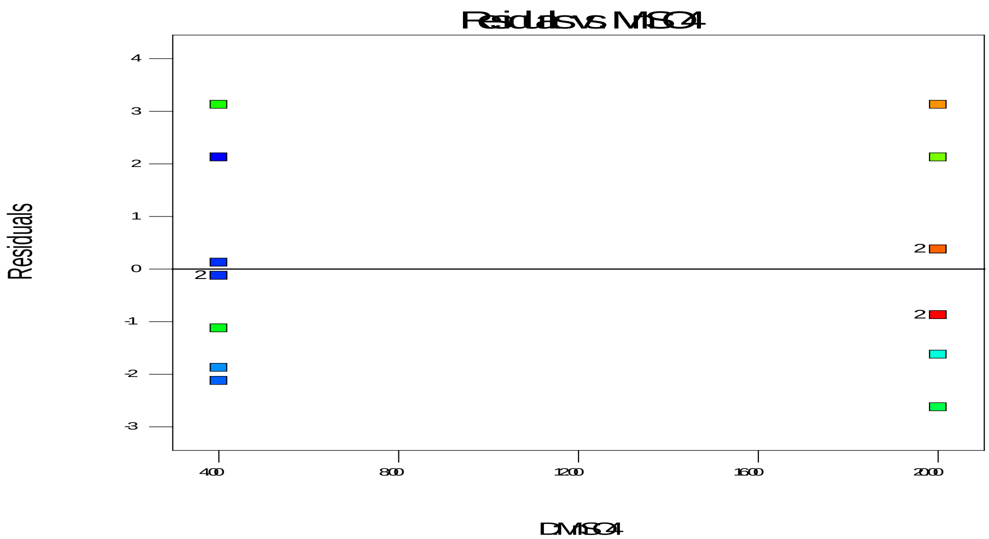

Normal Plot of Residuals
Normal % Probability
Residuals


Residuals vs. Predicted
Residuals
Predicted


Residuals vs Glucose
Residuals
4
3
2
1
0
-1
-2
-3
2000
2800
3600
4400
5200
6000
Glucose

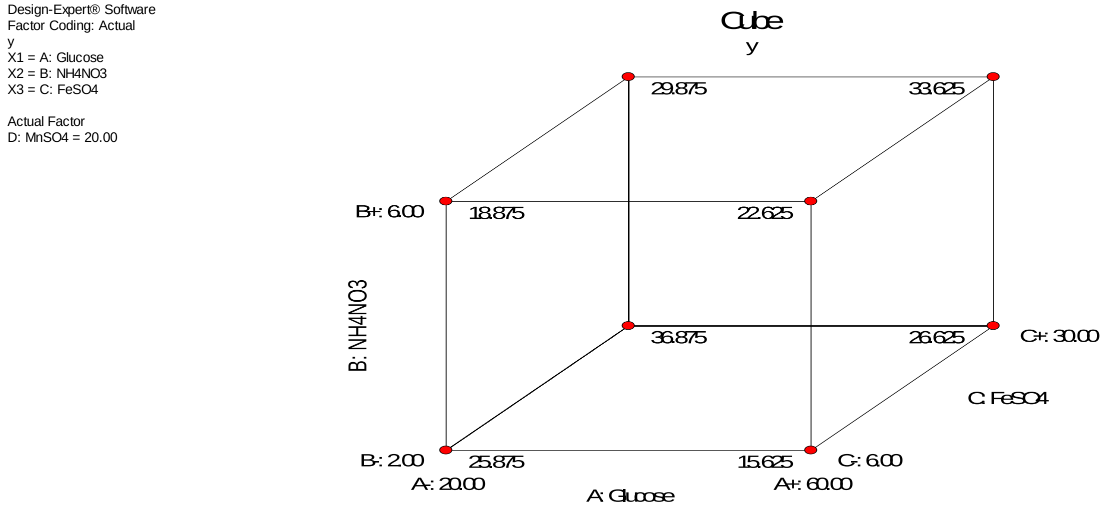

Residuals vs NH4NO3
Residuals
4
3
2
1
0
-1
-2
-3
200
300
400
500
600
B:NH4NO3





(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.



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:

			Fishbone		Hydroxyapatite	
Apatite	pH	Pb	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

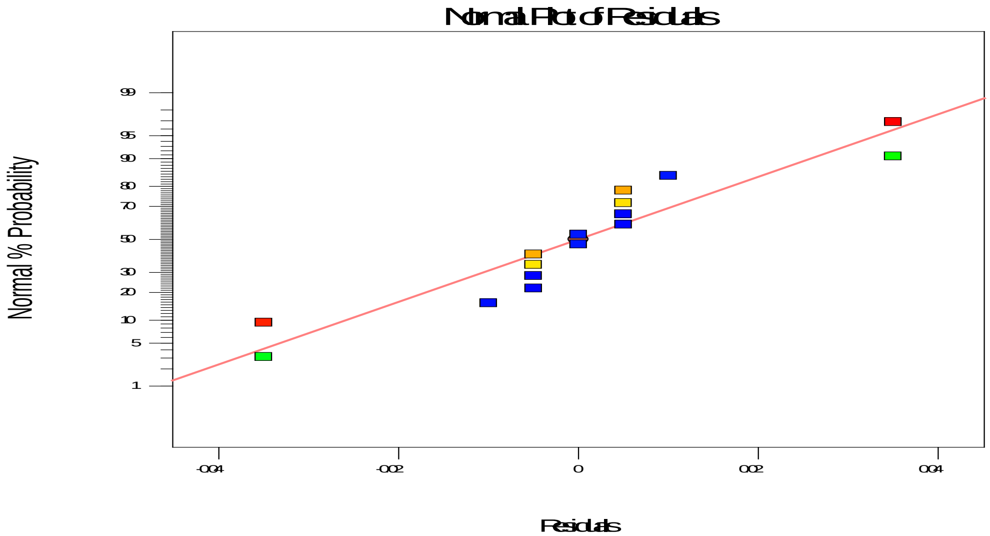
Response 1 Fishbone Pb
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

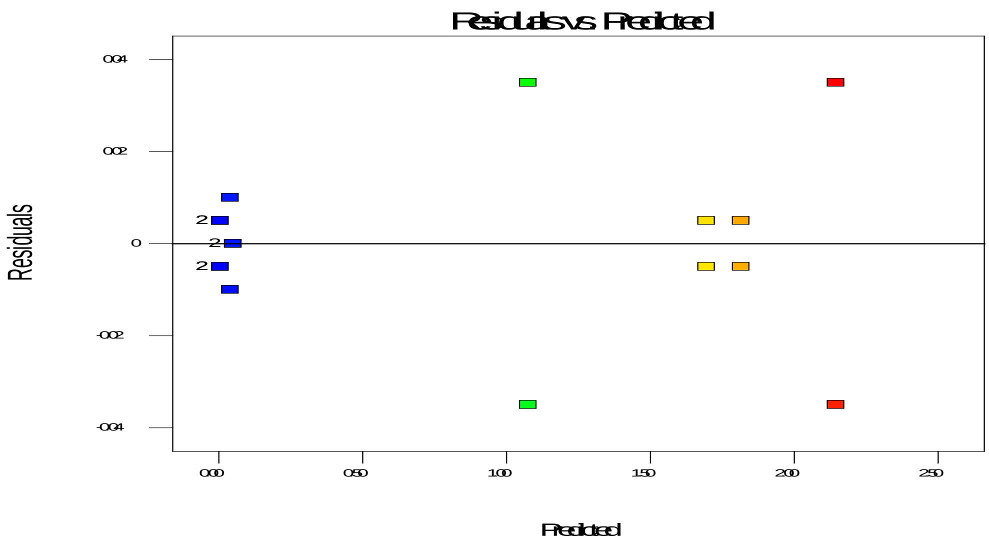
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	12.19	7	1.74	2629.41	< 0.0001	significant
A-Apatite	*0.27*	*1*	*0.27*	*400.34*	*< 0.0001*	
B-pH	*0.35*	*1*	*0.35*	*525.43*	*< 0.0001*	
C-Pb	*10.99*	*1*	*10.99*	*16587.51*	*< 0.0001*	
AB	*0.023*	*1*	*0.023*	*33.96*	*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				

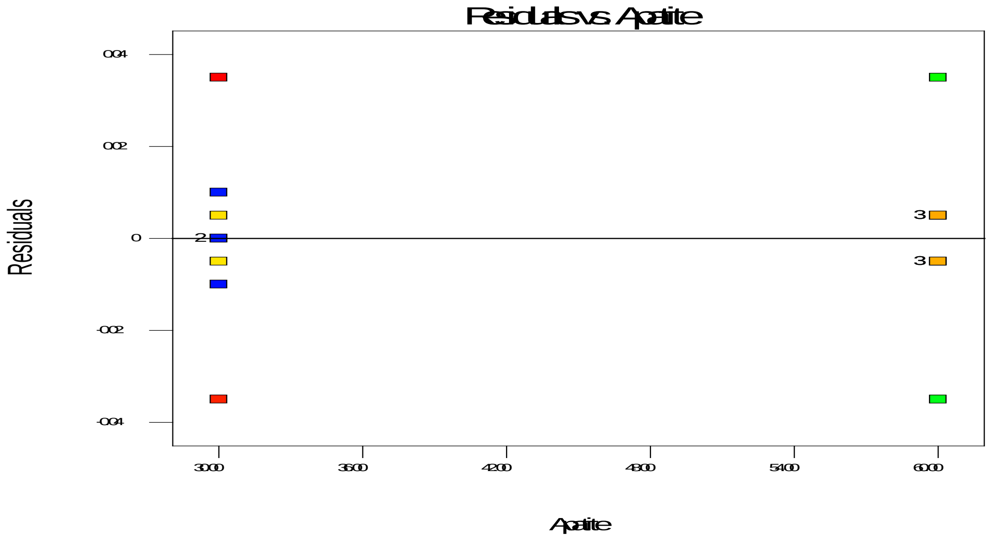
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.85	1	6.435E-003	0.84	0.87	
A-Apatite	-0.13	1	6.435E-003	-0.14	-0.11	1.00
B-pH	0.15	1	6.435E-003	0.13	0.16	1.00
C-Pb	0.83	1	6.435E-003	0.81	0.84	1.00
AB	0.038	1	6.435E-003	0.023	0.052	1.00
AC	-0.11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

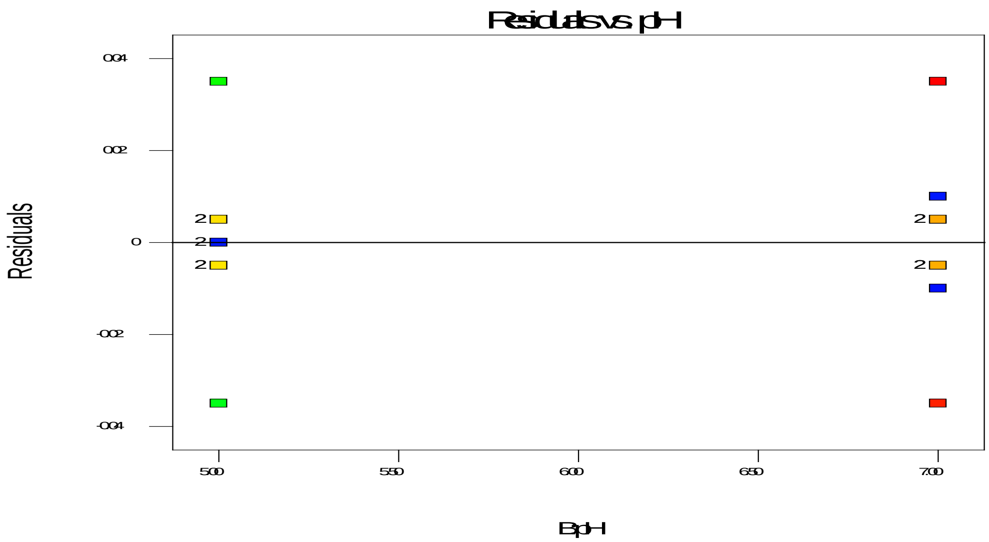
(b) Analyze the residuals from this response and comment on model adequacy.

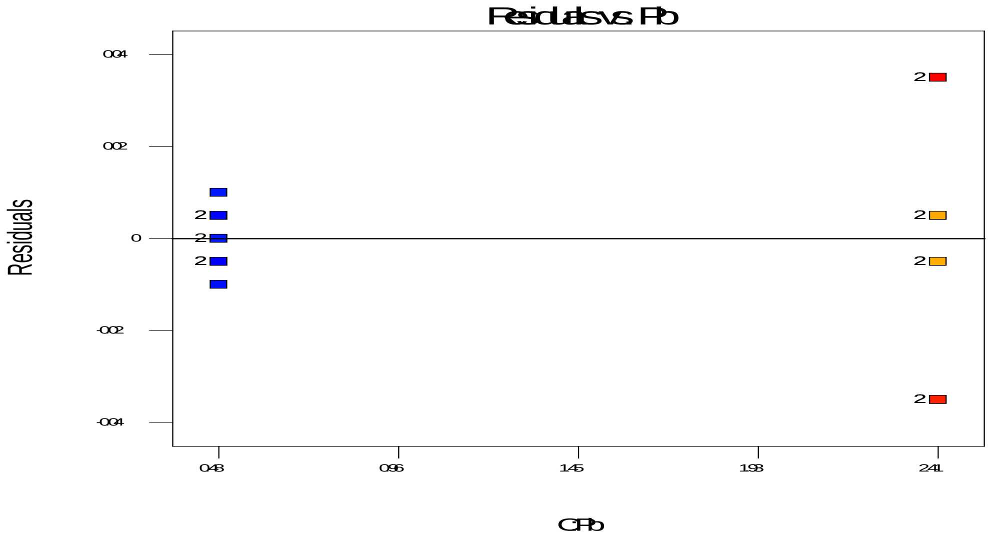
The normal plot identifies slightly thicker tails in the distribution of the residuals. The plots of residuals vs. predicted and residuals vs. the Pb effect identify nonconstant variance.






Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH



(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

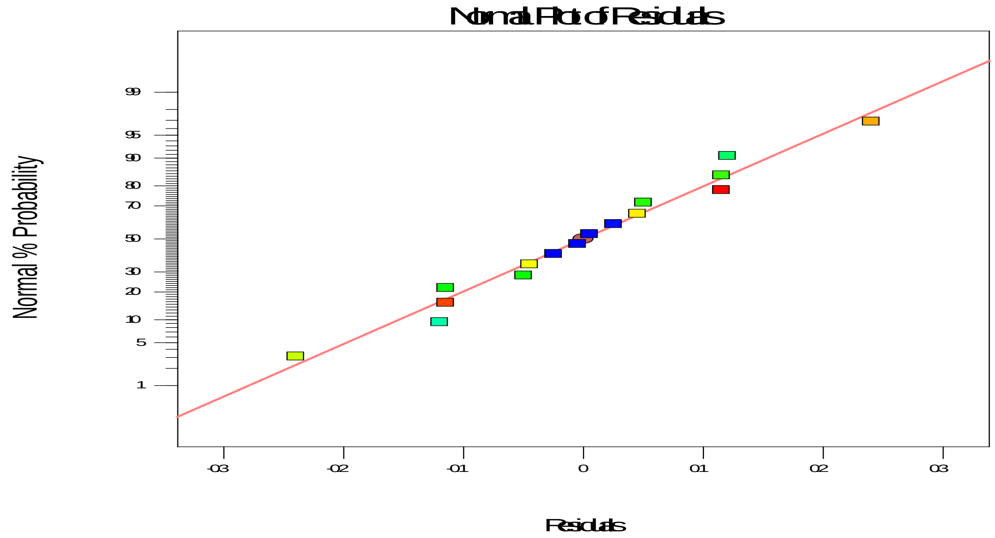
Response 1 Fishbone pH
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

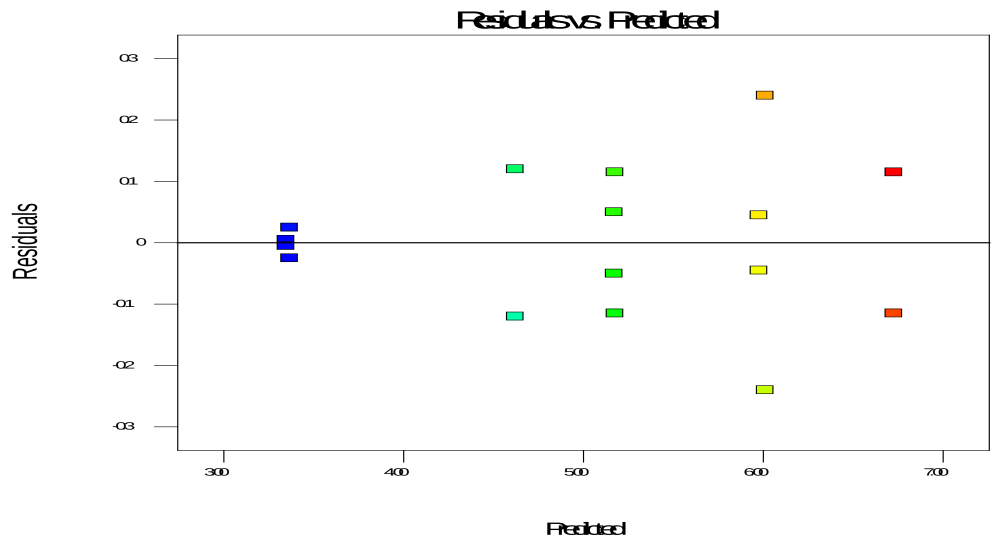
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	21.09	7	3.01	116.29	< 0.0001	significant
A-Apatite	*1.12*	*1*	*1.12*	*43.17*	*0.0002*	
B-pH	*8.14*	*1*	*8.14*	*314.08*	*< 0.0001*	
C-Pb	*9.84*	*1*	*9.84*	*379.98*	*< 0.0001*	
AB	*0.098*	*1*	*0.098*	*3.77*	*0.0881*	
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				

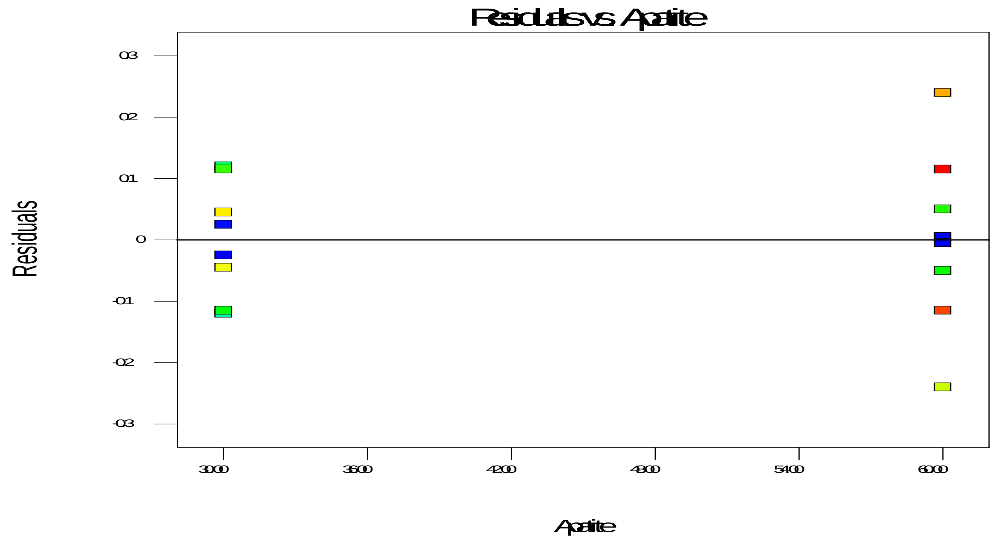
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	5.05	1	0.040	4.96	5.14	
A-Apatite	0.26	1	0.040	0.17	0.36	1.00
B-pH	0.71	1	0.040	0.62	0.81	1.00
C-Pb	-0.78	1	0.040	-0.88	-0.69	1.00
AB	-0.078	1	0.040	-0.17	0.015	1.00
AC	-0.27	1	0.040	-0.36	-0.18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

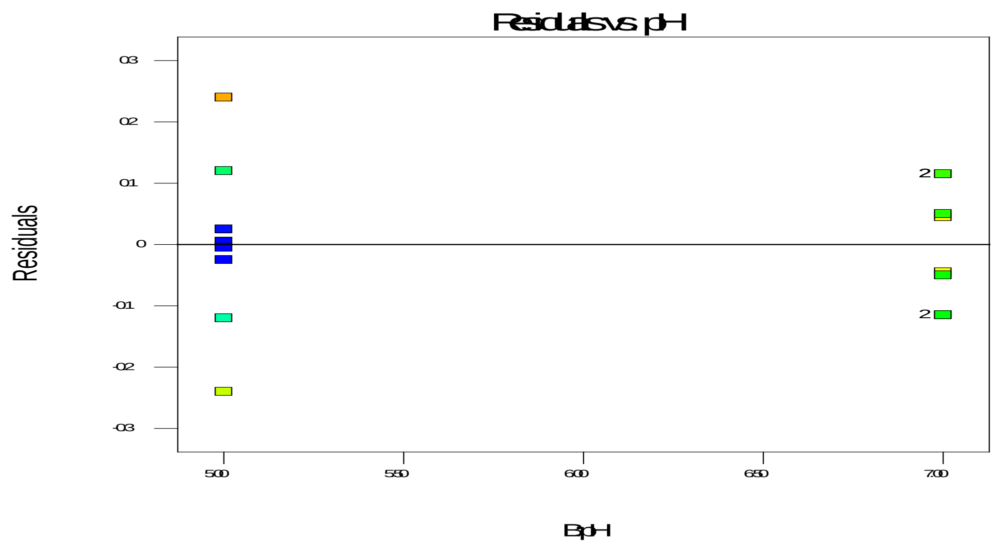
(d) Analyze the residuals from this response and comment on model adequacy.

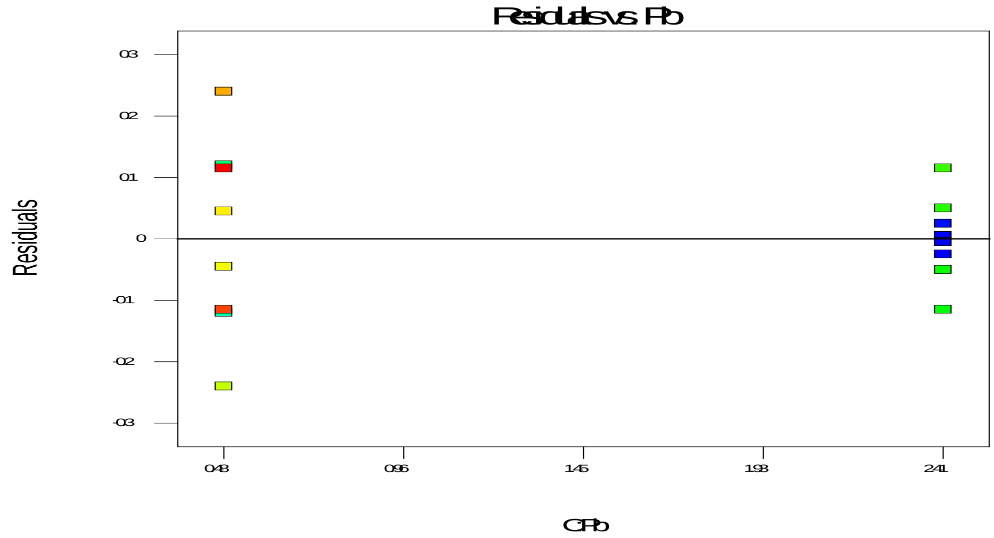
Although the normal probability plot is acceptable, the plots of residuals vs. predicted identifies nonconstant variance.


Normal Plot of Residuals
Normal % Probability
Residuals


Residuals vs. Predicted
Residuals
Predicted


Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH



(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

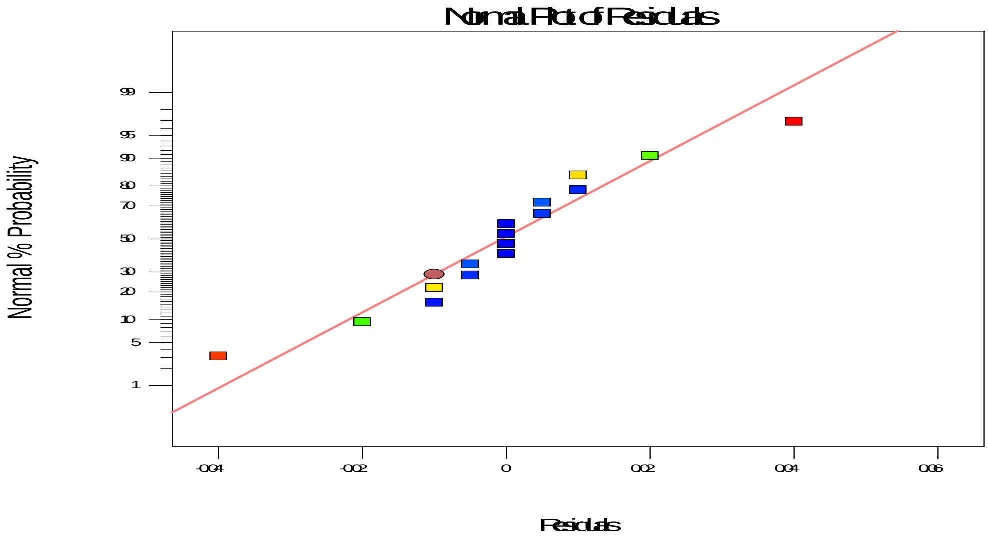
Response 1 Hydroxyapatite Pb
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

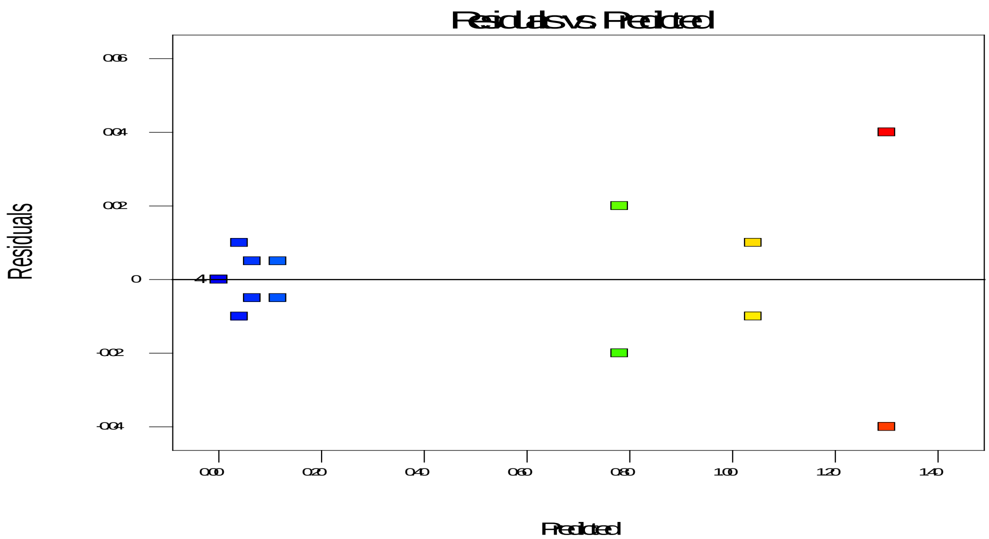
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	4.01	7	0.57	1018.21	< 0.0001	significant
A-Apatite	*0.54*	*1*	*0.54*	*960.40*	*< 0.0001*	
B-pH	*0.27*	*1*	*0.27*	*471.51*	*< 0.0001*	
C-Pb	*2.45*	*1*	*2.45*	*4354.18*	*< 0.0001*	
AB	*0.036*	*1*	*0.036*	*64.18*	*< 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				

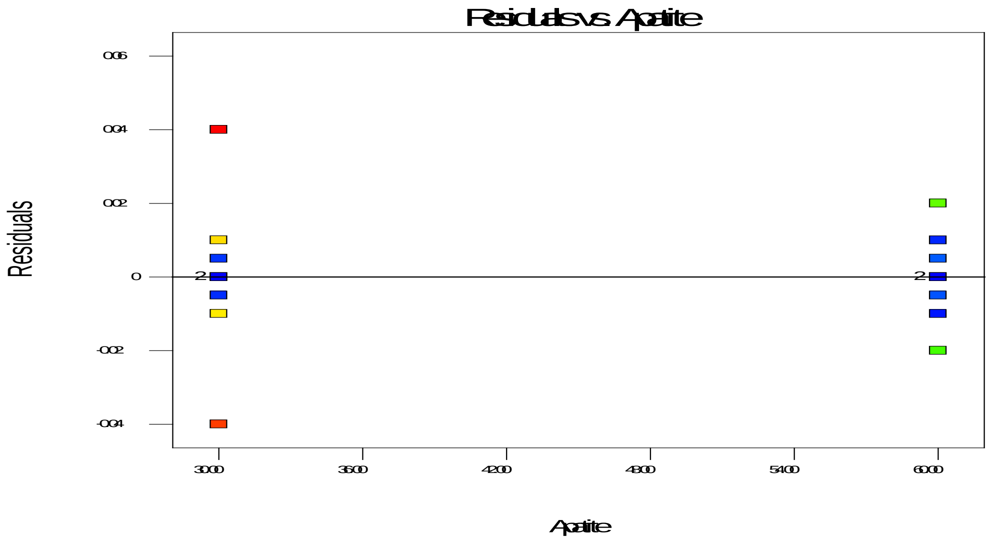
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.42	1	5.929E-003	0.40	0.43	
A-Apatite	-0.18	1	5.929E-003	-0.20	-0.17	1.00
B-pH	-0.13	1	5.929E-003	-0.14	-0.12	1.00
C-Pb	0.39	1	5.929E-003	0.38	0.40	1.00
AB	-0.048	1	5.929E-003	-0.061	-0.034	1.00
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

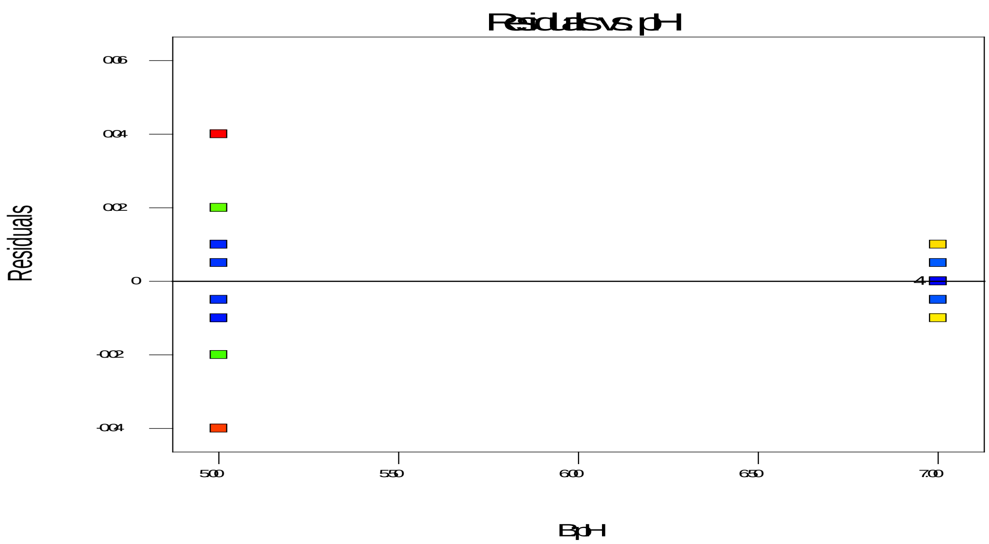
(f) Analyze the residuals from this response and comment on model adequacy.

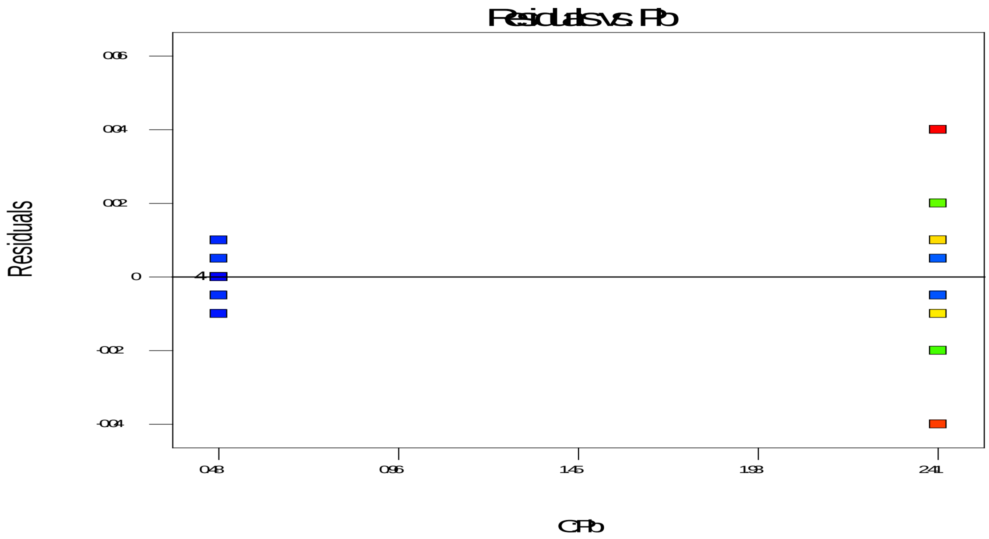
The normal plot identifies slightly thicker tails in the distribution of the residuals. The plots of residuals vs. predicted and residuals vs. the effects identifies nonconstant variance.


Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-0.04
-0.02
0
0.02
0.04
0.06
Residuals


Residuals vs. Predicted
Residuals
0.06
0.04
0.02
0
-0.02
-0.04
0.00
0.20
0.40
0.60
0.80
1.00
1.20
1.40
Predicted


Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH



(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

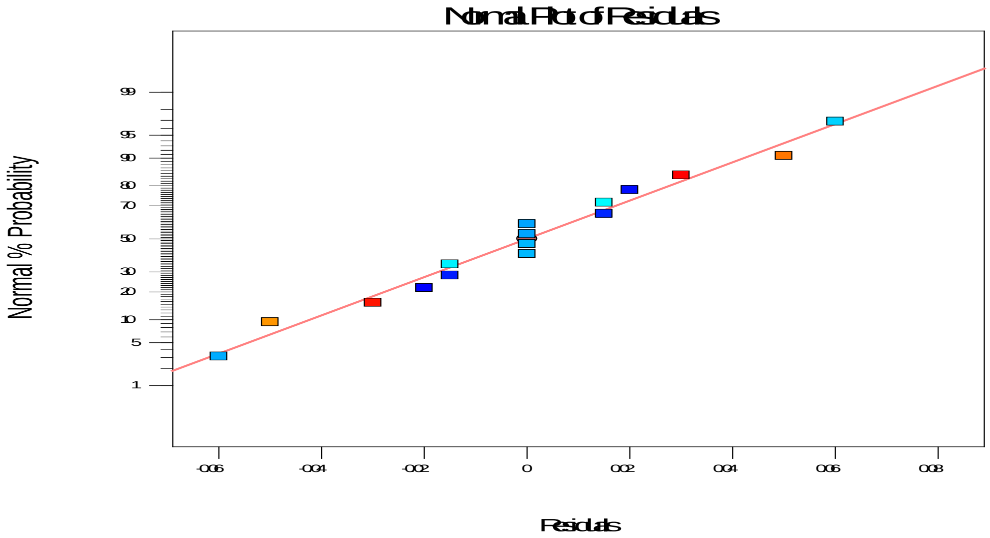
Response 1 Hydroxyapatite pH
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

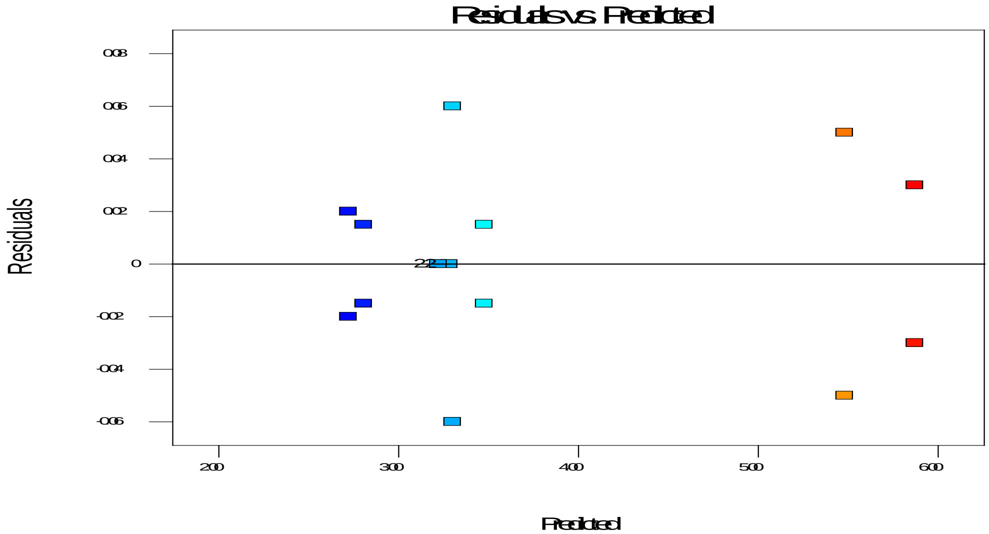
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	20.44	7	2.92	1487.66	< 0.0001	significant
A-Apatite	*0.084*	*1*	*0.084*	*42.85*	*0.0002*	
B-pH	*8.82*	*1*	*8.82*	*4494.73*	*< 0.0001*	
C-Pb	*8.15*	*1*	*8.15*	*4153.39*	*< 0.0001*	
AB	*0.13*	*1*	*0.13*	*64.22*	*< 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				

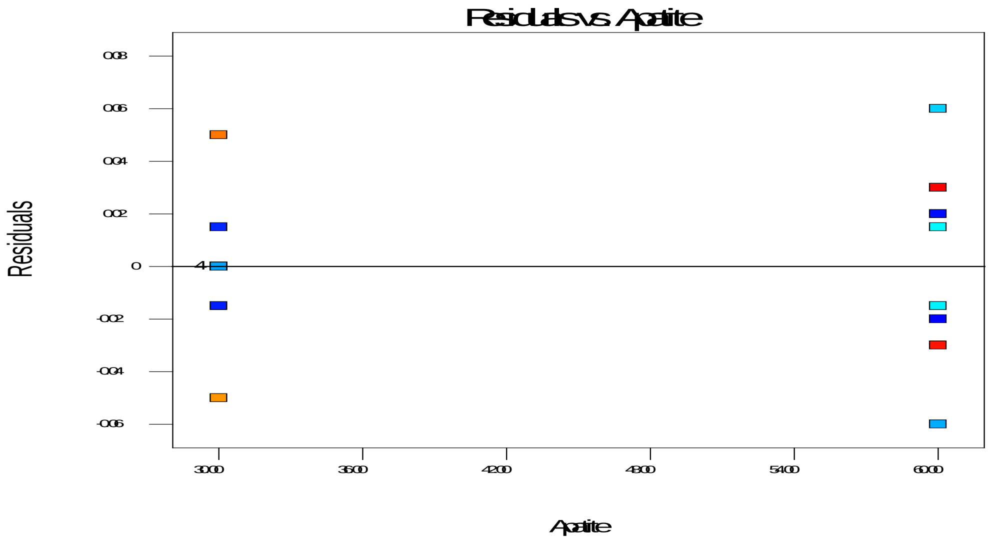
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	3.77	1	0.011	3.74	3.79	
A-Apatite	0.072	1	0.011	0.047	0.098	1.00
B-pH	0.74	1	0.011	0.72	0.77	1.00
C-Pb	-0.71	1	0.011	-0.74	-0.69	1.00
AB	0.089	1	0.011	0.063	0.11	1.00
AC	-0.030	1	0.011	-0.056	-4.461E-003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

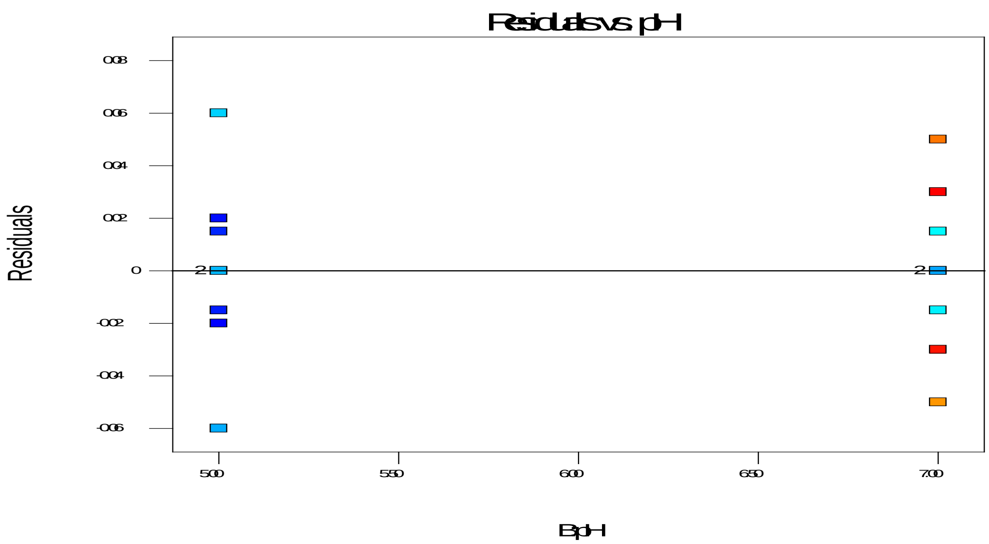
(h) Analyze the residuals from this response and comment on model adequacy.

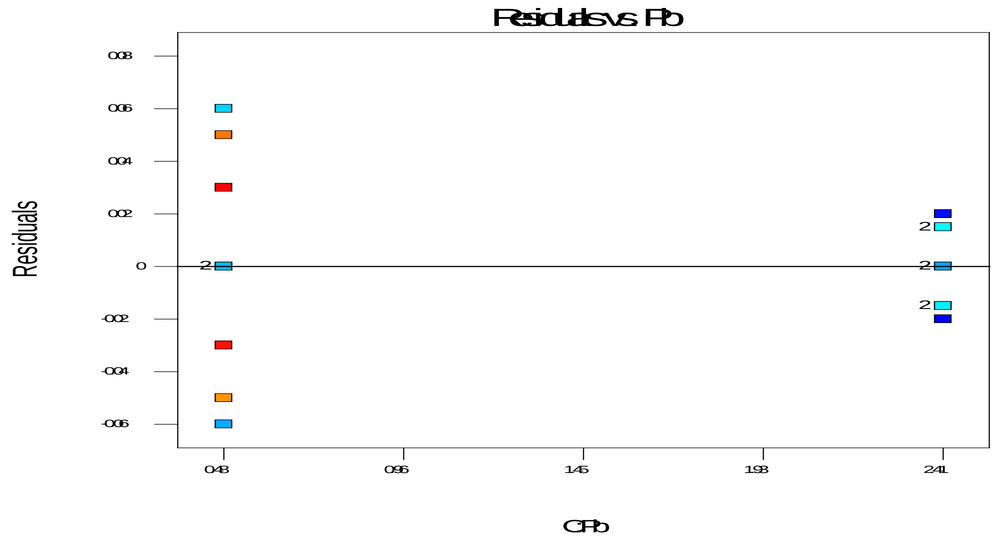
The only potential concern with the residual plots is the nonconstant variance shown in the plot of residuals vs. pH.


Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-0.06
-0.04
-0.02
0
0.02
0.04
0.06
0.08


Residuals
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
200
300
400
500
600


Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals



(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

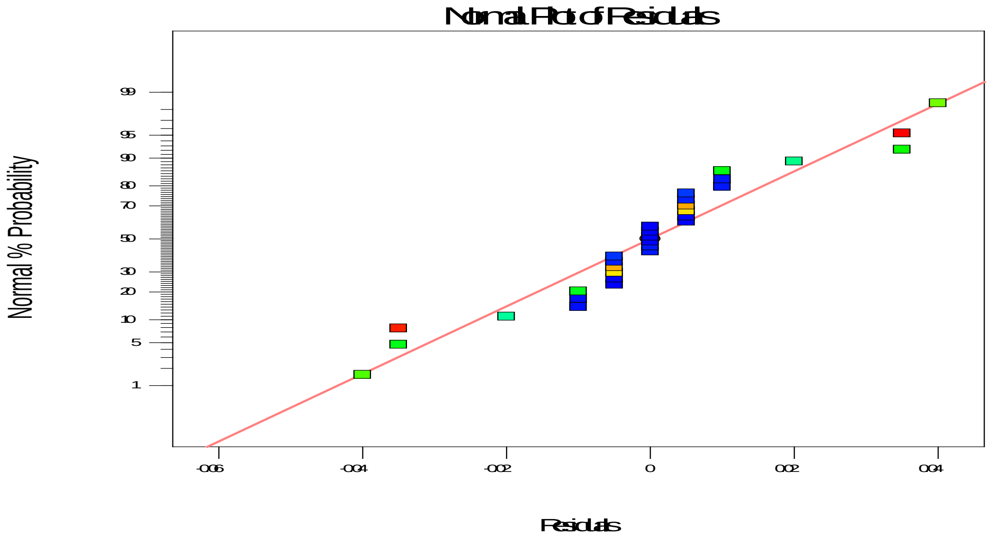
The ANOVA below identifies factor *D*, the type of apatite, as being very significant.

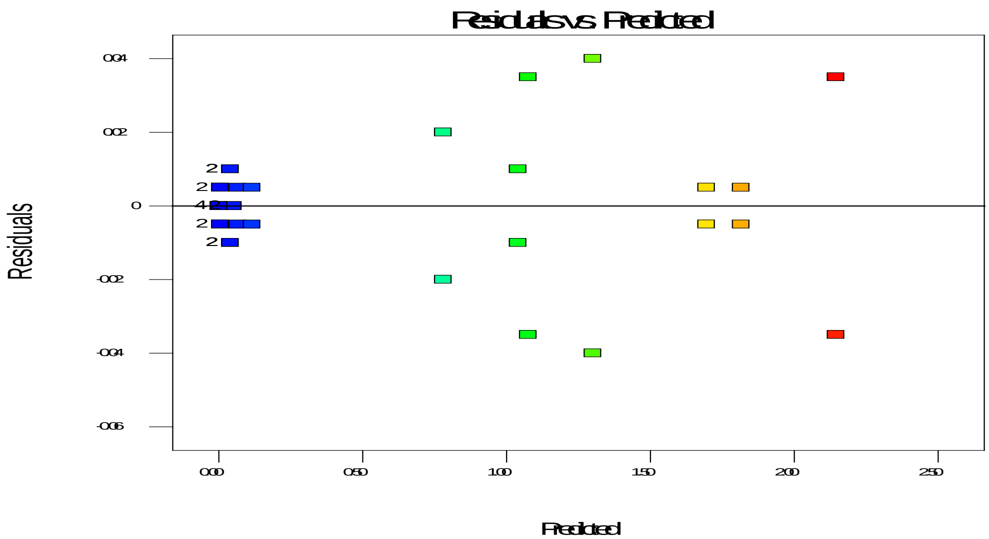
Design Expert Output

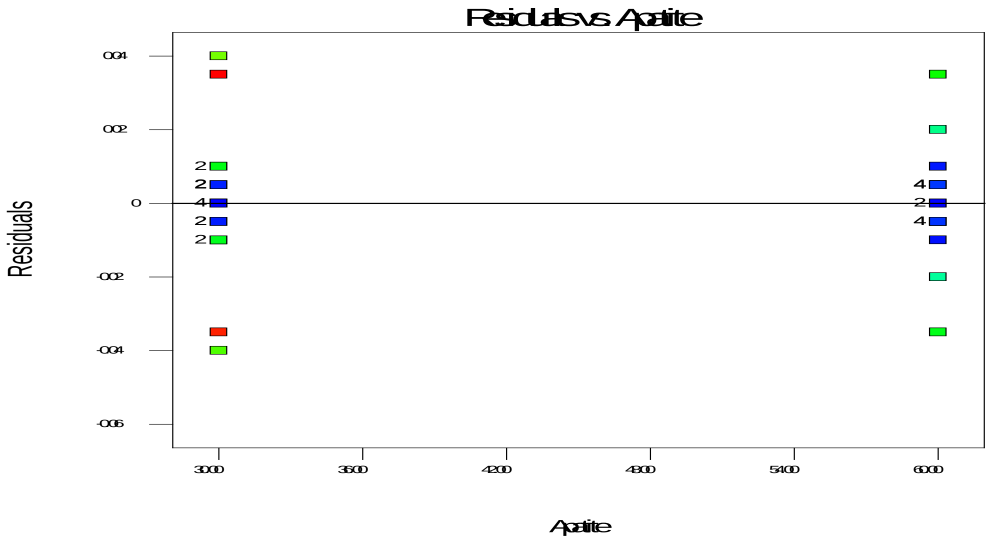
Response 1 Pb Response
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

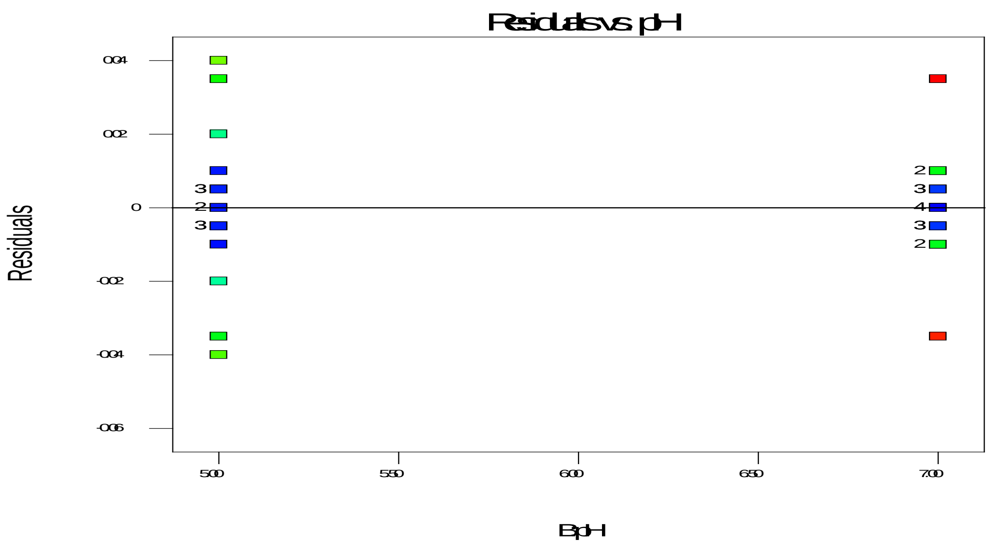
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	17.73	15	1.18	1929.32	< 0.0001	significant
A-Apatite	*0.78*	*1*	*0.78*	*1275.51*	*< 0.0001*	
B-pH	*2.813E-003*	*1*	*2.813E-003*	*4.59*	*0.0478*	
C-Pb	*11.91*	*1*	*11.91*	*19440.33*	*< 0.0001*	
D-Type	*1.52*	*1*	*1.52*	*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				

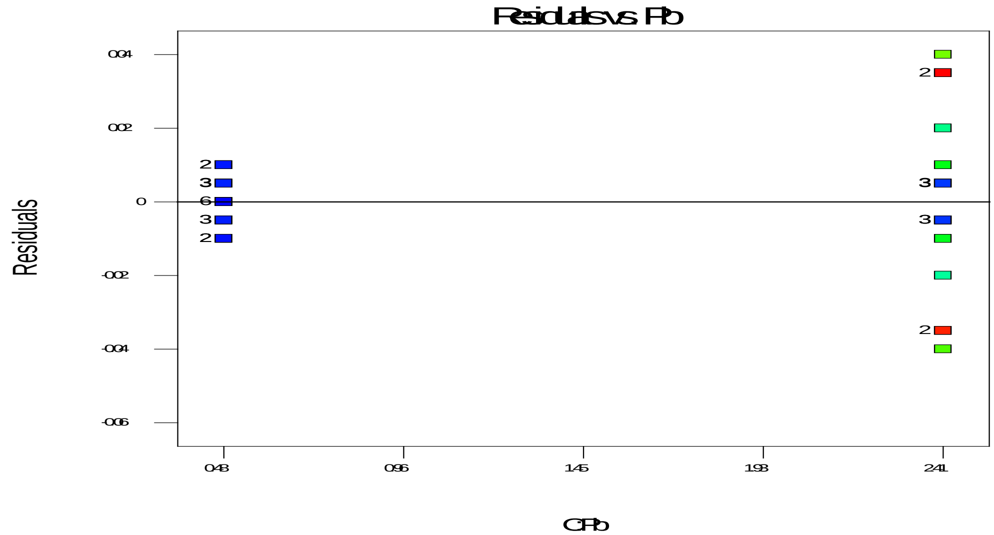
The residual plots below identify concerns, so a power transformation with lambda of 0.7 was applied to the Pb response.

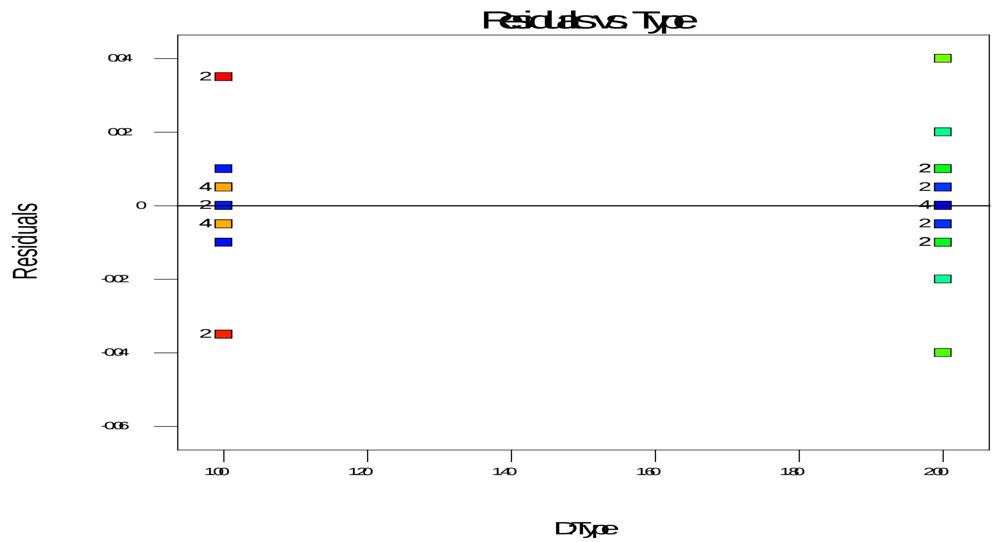

Normal Plot of Residuals
Normal % Probability
Residuals


Residuals vs. Predicted
Residuals
Predicted


Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH





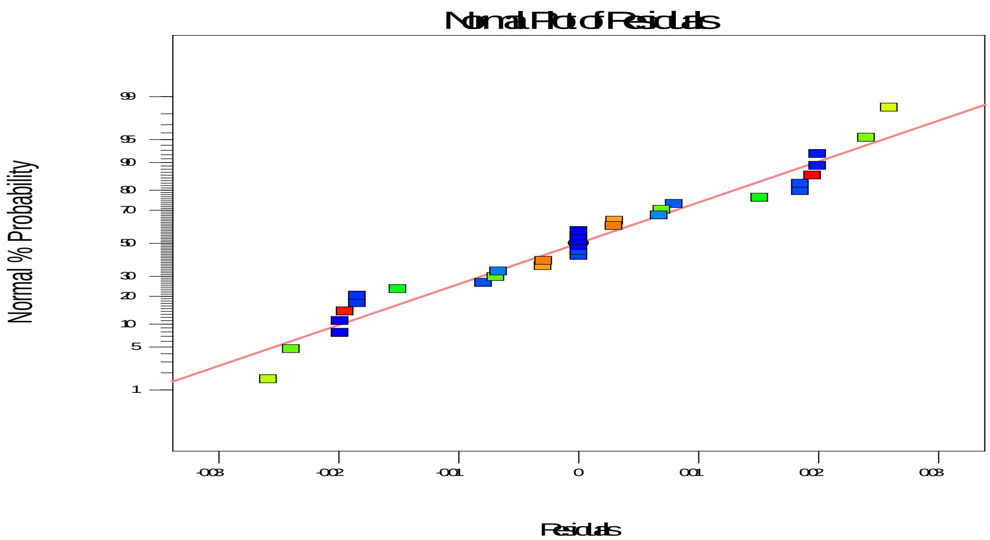
Design Expert Output

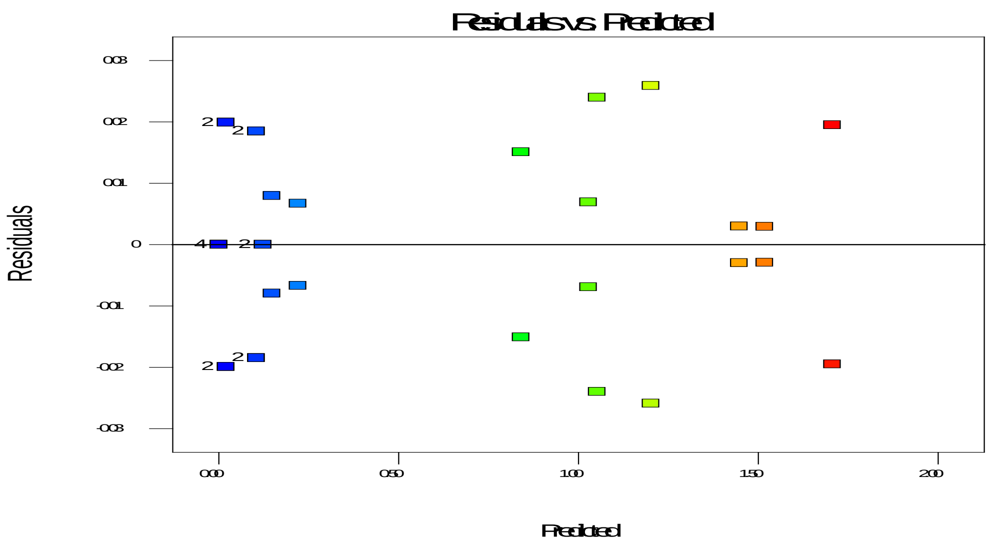
Response 1 Pb Response
Transform: Power Lambda: 0.7 Constant: 0
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

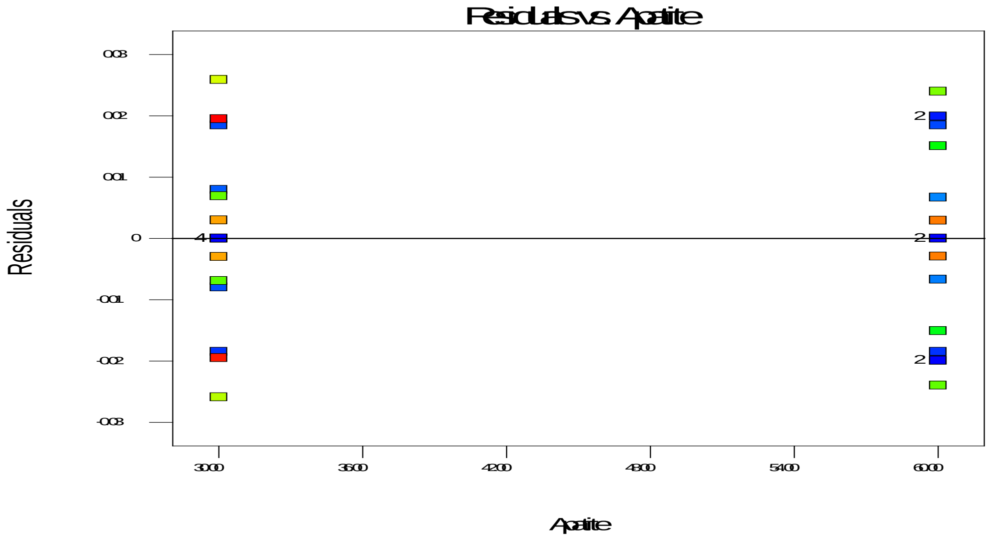
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	12.10	15	0.81	1844.15	< 0.0001	significant
A-Apatite	*0.49*	*1*	*0.49*	*1123.40*	*< 0.0001*	
B-pH	*0.014*	*1*	*0.014*	*32.86*	*< 0.0001*	
C-Pb	*9.02*	*1*	*9.02*	*20612.51*	*< 0.0001*	
D-Type	*0.75*	*1*	*0.75*	*1712.43*	*< 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*	

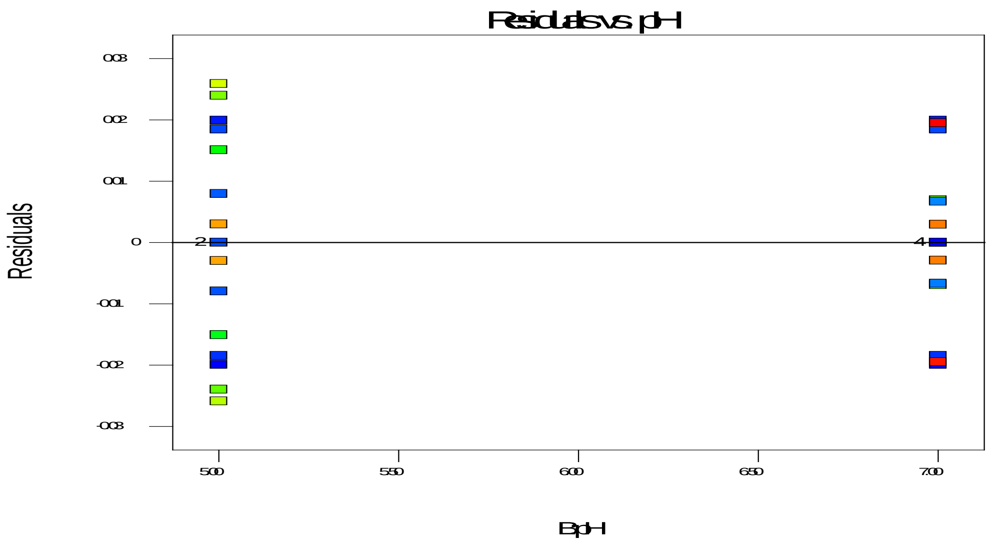
ABD	*0.049*	*1*	*0.049*	*112.97*	*< 0.0001*
ACD	*0.067*	*1*	*0.067*	*152.27*	*< 0.0001*
BCD	*0.21*	*1*	*0.21*	*472.40*	*< 0.0001*
ABCD	*0.057*	*1*	*0.057*	*131.40*	*< 0.0001*
Pure Error	6.999E-003	16	4.374E-004		
Cor Total	12.11	31			

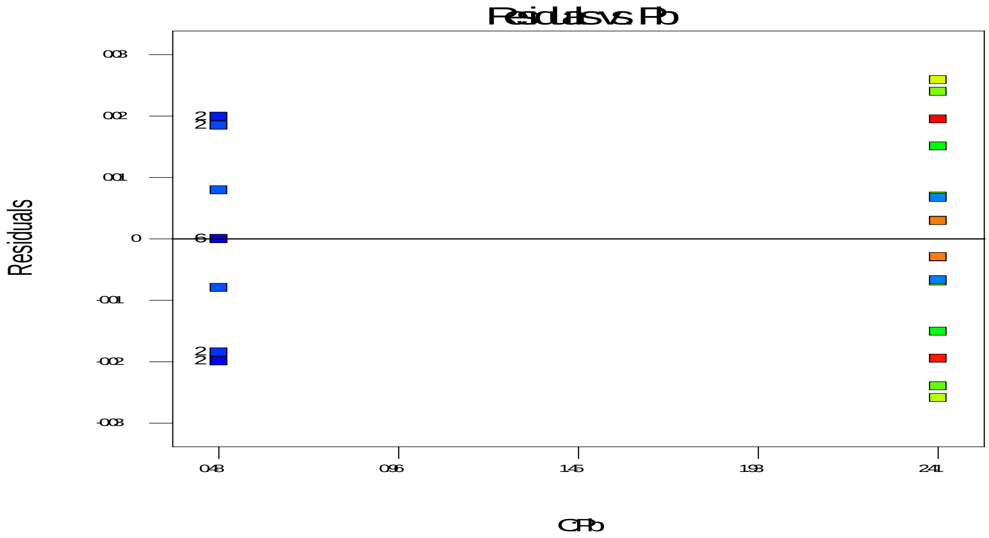
There are no concerns with the residual plots shown below.

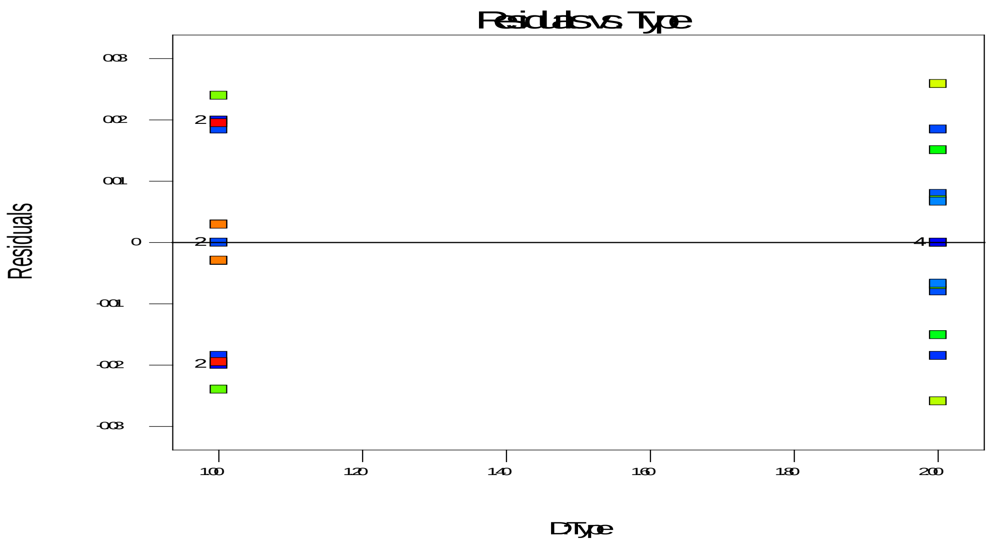





Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH





The analysis for the pH response is shown below.

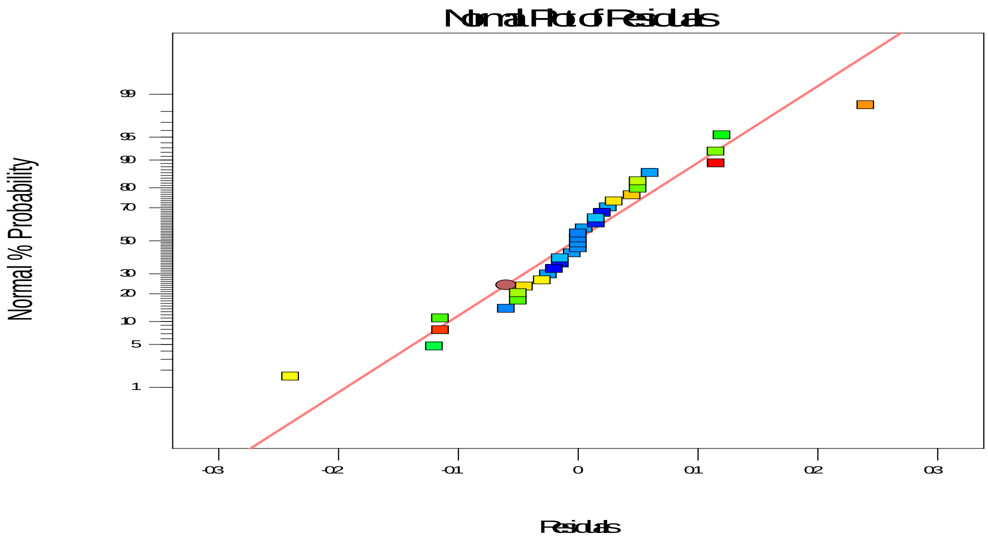
Design Expert Output

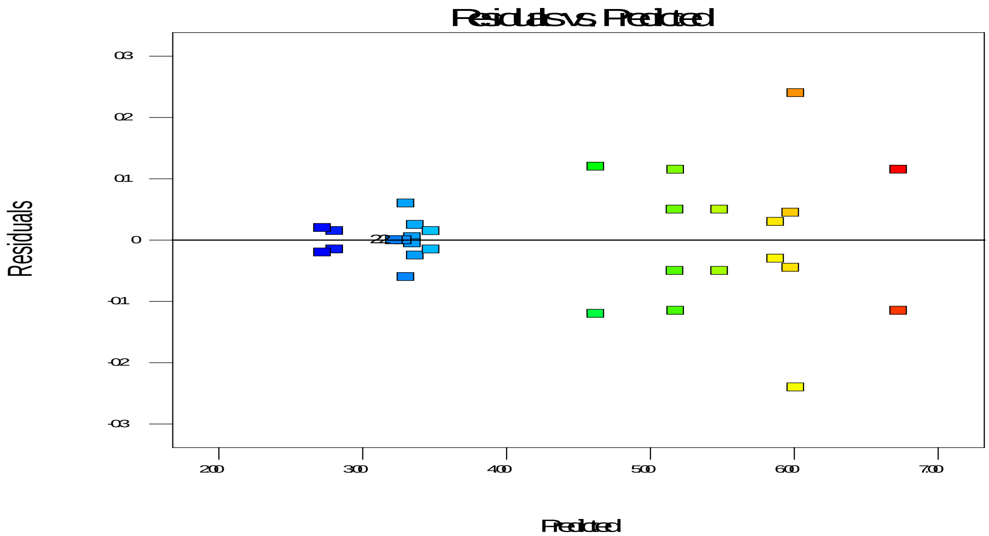
Response 2 pH Response
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

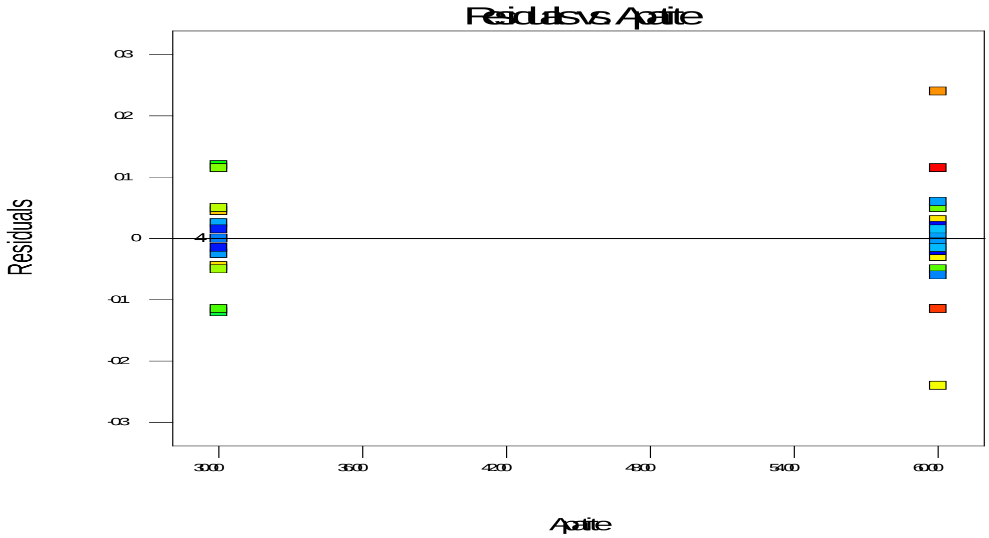
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	54.62	15	3.64	261.32	< 0.0001	significant
A-Apatite	*0.91*	*1*	*0.91*	*65.15*	*< 0.0001*	
B-pH	*16.95*	*1*	*16.95*	*1216.47*	*< 0.0001*	
C-Pb	*17.96*	*1*	*17.96*	*1288.54*	*< 0.0001*	
D-Type	*13.09*	*1*	*13.09*	*939.72*	*< 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*	

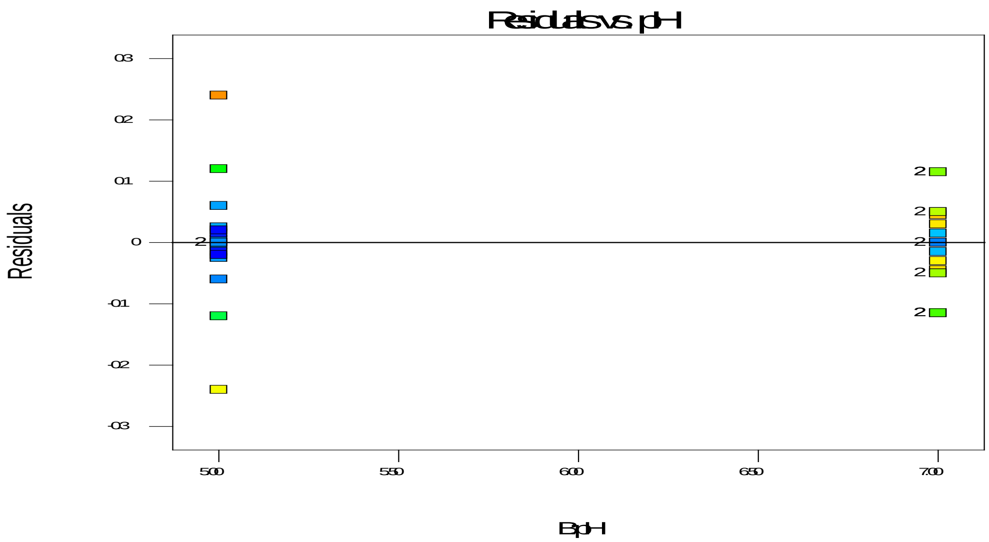
ABC	*0.049*	*1*	*0.049*	*3.50*	*0.0796*
ABD	*0.22*	*1*	*0.22*	*15.99*	*0.0010*
ACD	*0.46*	*1*	*0.46*	*33.24*	*< 0.0001*
BCD	*3.33*	*1*	*3.33*	*239.31*	*< 0.0001*
ABCD	*0.059*	*1*	*0.059*	*4.21*	*0.0569*
Pure Error	0.22	16	0.014		
Cor Total	54.84	31			

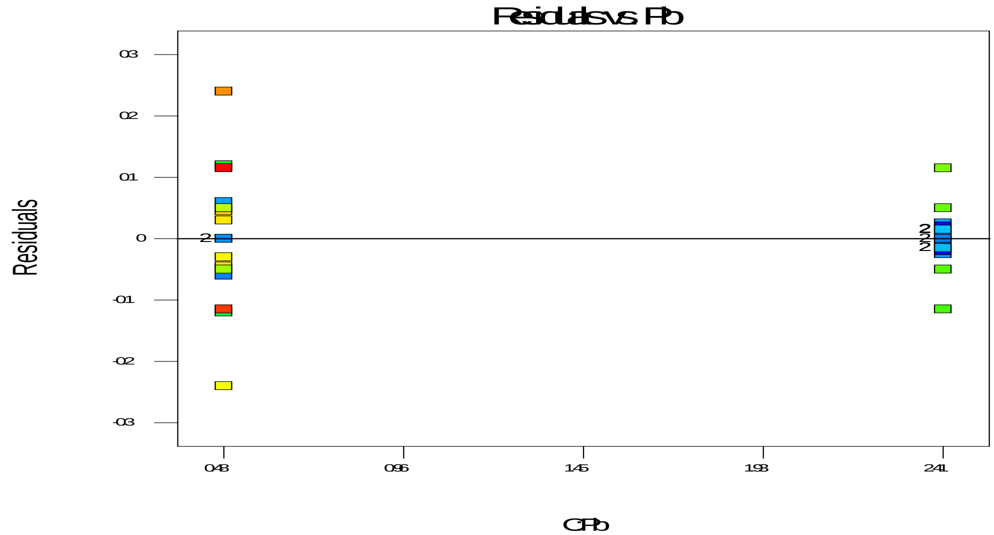
Nonconstant variance is identified in the residual plots below, so an inverse square root transformation was applied.

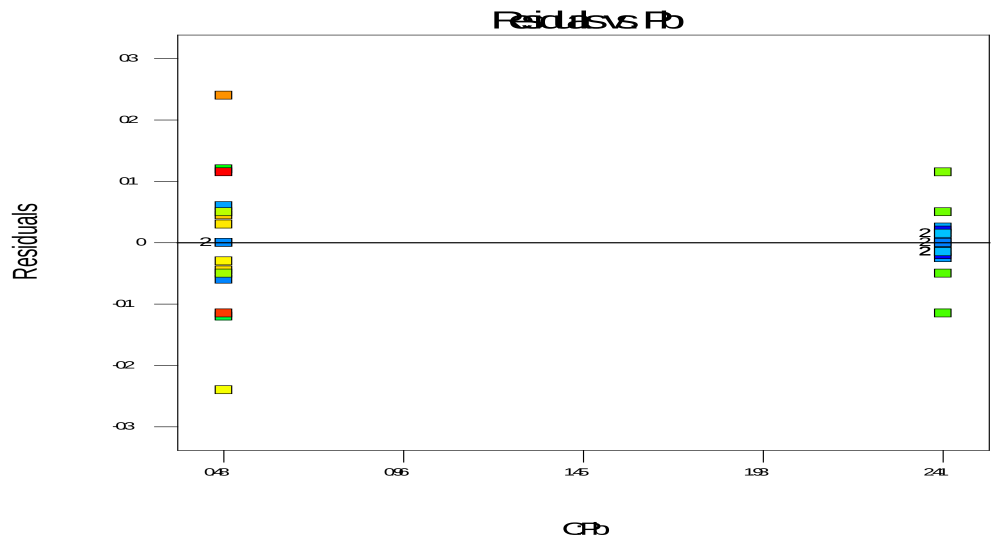





Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH





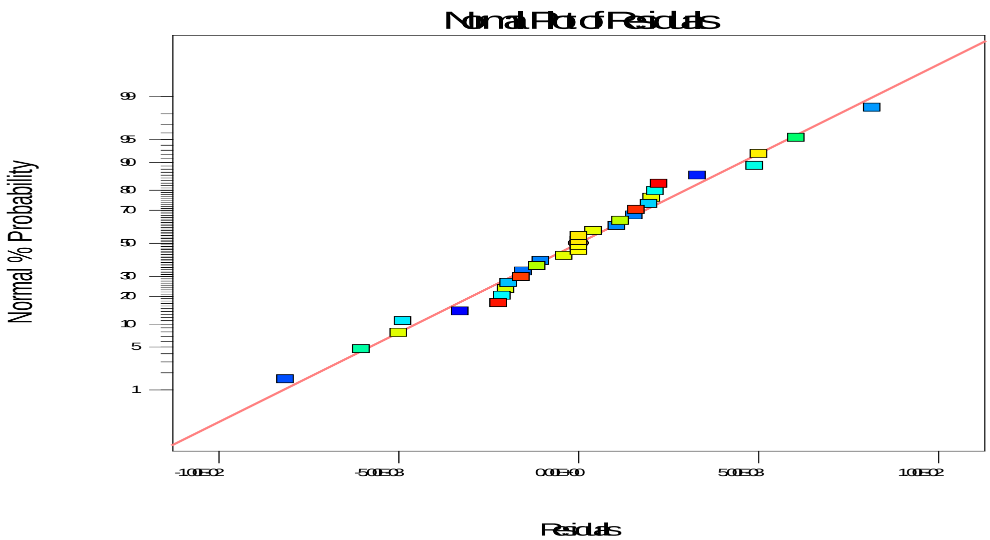
Design Expert Output

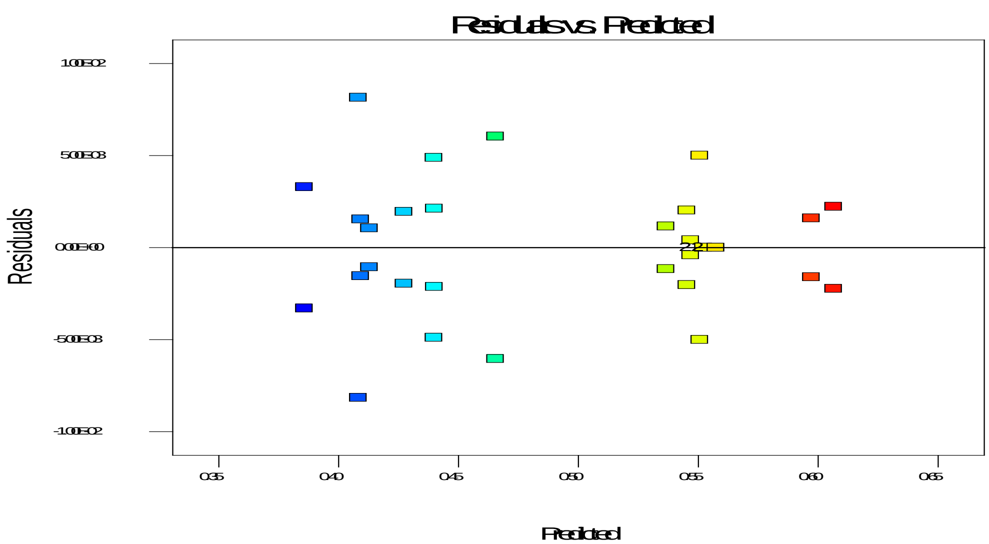
Response **2** **pH Response**
Transform: **Inverse Sqrt** **Constant:** **0**
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

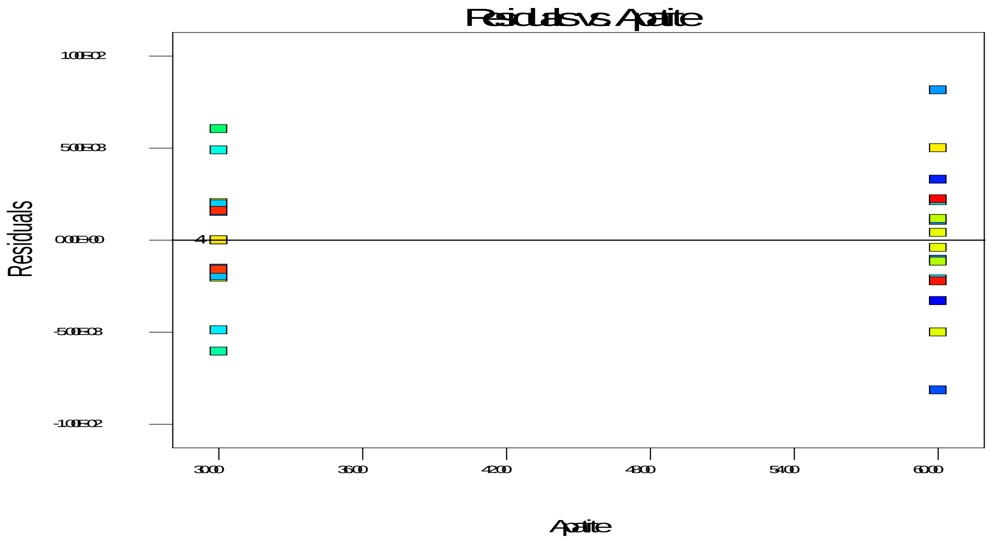
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	0.17	15	0.011	483.43	< 0.0001	significant
A-Apatite	*1.418E-003*	*1*	*1.418E-003*	*60.44*	*< 0.0001*	
B-pH	*0.055*	*1*	*0.055*	*2346.06*	*< 0.0001*	
C-Pb	*0.054*	*1*	*0.054*	*2303.65*	*< 0.0001*	
D-Type	*0.045*	*1*	*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*	

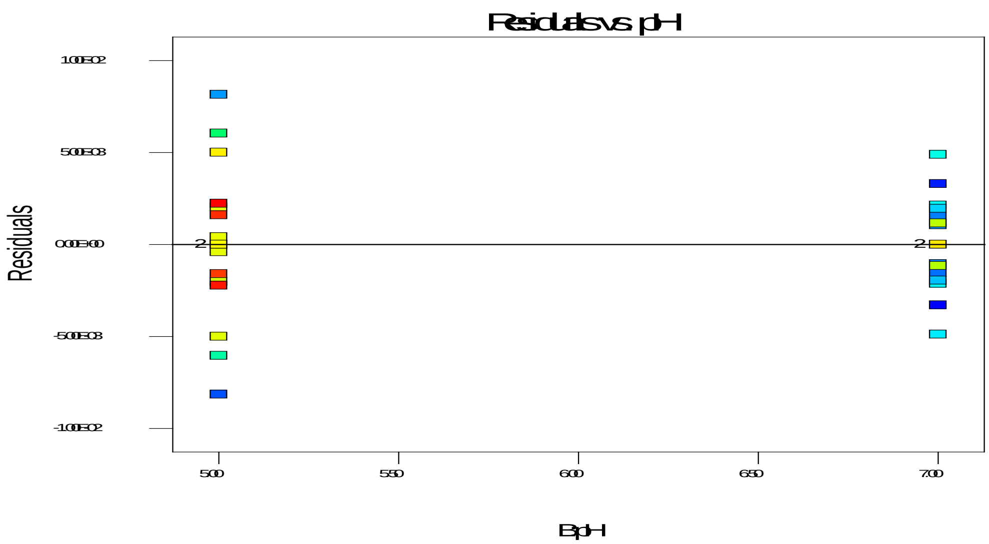
ABD	*7.072E-004*	*1*	*7.072E-004*	*30.14*	*< 0.0001*
ACD	*7.593E-004*	*1*	*7.593E-004*	*32.36*	*< 0.0001*
BCD	*0.010*	*1*	*0.010*	*437.95*	*< 0.0001*
ABCD	*4.034E-005*	*1*	*4.034E-005*	*1.72*	*0.2083*
Pure Error	3.755E-004	16	2.347E-005		
Cor Total	0.17	31			

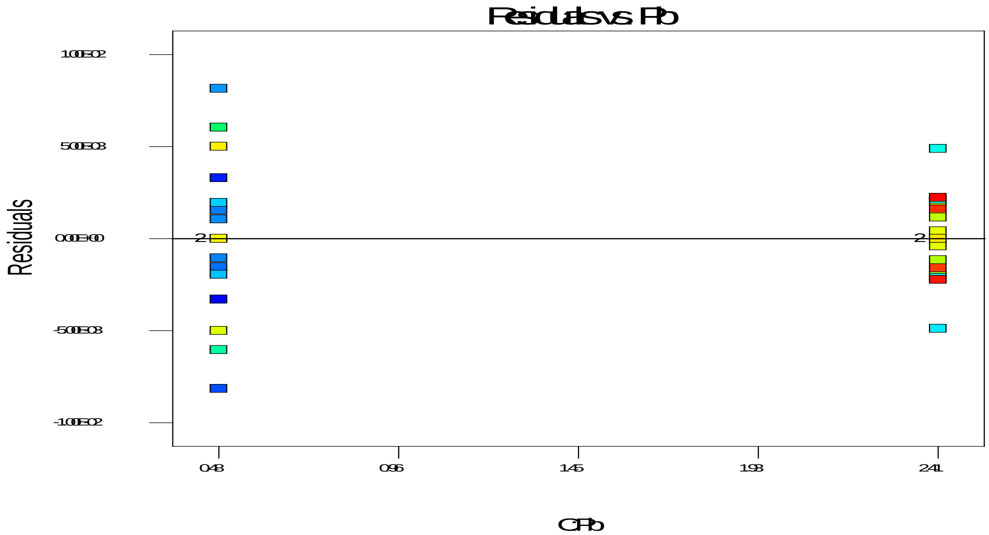
There are no concerns with the residual plots shown below.

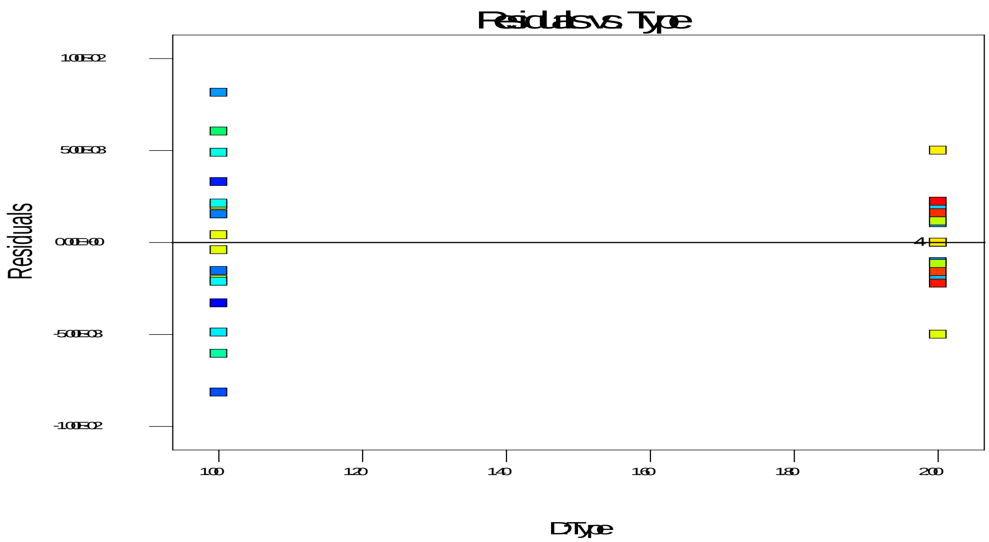





Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals
B:pH





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 $x_1, x_2, \ldots, x_k$ 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 $\dfrac{\sigma^2}{n2^k}$ and that the covariance between any pair of regression coefficients is zero.

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

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

where $\mathbf{x}' = [x_1, x_2, ..., x_k]$ 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 $x_1, x_2, ..., x_p$ potentially include interaction terms among the original *k* variables. Now the variance of the predicted response is

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

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-α)% confidence interval on the true mean response at the point $x_1, x_2, \ldots, x_k$ in the design space.

The confidence interval is

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

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 Hierarchial 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			
Lack of Fit	*93.00*	*3*	*31.00*	*1.03*	*0.4067*	*not significant*
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]

Source	**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*(Effect) and 100(1-α) 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

Chapter 7
Blocking and Confounding in the 2^k Factorial Design
Solutions

7.1 Consider the experiment described in Problem 6.1. Analyze this experiment assuming that each replicate represents a block of a single production shift.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Cutting Speed (*A*)	0.67	1	0.67	<1
Tool Geometry (*B*)	770.67	1	770.67	22.38*
Cutting Angle (*C*)	280.17	1	280.17	8.14*
AB	16.67	1	16.67	<1
AC	468.17	1	468.17	13.60*
BC	48.17	1	48.17	1.40
ABC	28.17	1	28.17	<1
Blocks	0.58	2	0.29	
Error	482.08	14	34.43	
Total	2095.33	23		

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**	
Block	0.58	2	0.29			
Model	1519.67	4	379.92	11.23	0.0001	significant
A	*0.67*	*1*	*0.67*	*0.020*	*0.8900*	
B	*770.67*	*1*	*770.67*	*22.78*	*0.0002*	
C	*280.17*	*1*	*280.17*	*8.28*	*0.0104*	
AC	*468.17*	*1*	*468.17*	*13.84*	*0.0017*	
Residual	575.08	17	33.83			
Cor Total	2095.33	23				

The Model F-value of 11.23 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.

These results agree with the results from Problem 6.1. Tool geometry, cutting angle and the interaction between cutting speed and cutting angle are significant at the 5% level. The Design Expert program also includes factor *A*, cutting speed, in the model to preserve hierarchy.

7.2. Consider the experiment described in Problem 6.5. Analyze this experiment assuming that each one of the four replicates represents a block.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Bit Size (*A*)	1107.23	1	1107.23	364.22*
Cutting Speed (*B*)	227.26	1	227.26	74.76*
AB	303.63	1	303.63	99.88*
Blocks	44.36	3	14.79	
Error	27.36	9	3.04	
Total	1709.83	15		

These results agree with those from Problem 6.5. Bit size, cutting speed and their interaction are significant at the 1% level.

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	
Block	44.36	3	14.79			
Model	1638.11	3	546.04	179.61	< 0.0001	significant
A	*1107.23*	*1*	*1107.23*	*364.21*	*< 0.0001*	
B	*227.26*	*1*	*227.26*	*74.75*	*< 0.0001*	
AB	*303.63*	*1*	*303.63*	*99.88*	*< 0.0001*	
Residual	27.36	9	3.04			
Cor Total	1709.83	15				

The Model F-value of 179.61 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.

7.3. Consider the alloy cracking experiment described in Problem 6.15. Suppose that only 16 runs could be made on a single day, so each replicate was treated as a block. Analyze the experiment and draw conclusions.

The analysis of variance for the full model is as follows:

Design Expert Output

Response: Crack Lengthin 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	
Block	0.016	1	0.016			
Model	570.95	15	38.06	445.11	< 0.0001	significant
A	*72.91*	*1*	*72.91*	*852.59*	*< 0.0001*	
B	*126.46*	*1*	*126.46*	*1478.83*	*< 0.0001*	
C	*103.46*	*1*	*103.46*	*1209.91*	*< 0.0001*	
D	*30.66*	*1*	*30.66*	*358.56*	*< 0.0001*	
AB	*29.93*	*1*	*29.93*	*349.96*	*< 0.0001*	
AC	*128.50*	*1*	*128.50*	*1502.63*	*< 0.0001*	
AD	*0.047*	*1*	*0.047*	*0.55*	*0.4708*	
BC	*0.074*	*1*	*0.074*	*0.86*	*0.3678*	
BD	*0.018*	*1*	*0.018*	*0.21*	*0.6542*	
CD	*0.047*	*1*	*0.047*	*0.55*	*0.4686*	
ABC	*78.75*	*1*	*78.75*	*920.92*	*< 0.0001*	

ABD	*0.077*	*1*	*0.077*	*0.90*	*0.3582*
ACD	*2.926E-003*	*1*	*2.926E-003*	*0.034*	*0.8557*
BCD	*0.010*	*1*	*0.010*	*0.12*	*0.7352*
ABCD	*1.596E-003*	*1*	*1.596E-003*	*0.019*	*0.8931*
Residual	1.28	15	0.086		
Cor Total	572.25	31			

The Model F-value of 445.11 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.

The analysis of variance for the reduced model based on the significant factors is shown below. The BC interaction was included to preserve hierarchy.

Design Expert Output

Response: Crack Lengthin 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**	
Block	0.016	1	0.016			
Model	570.74	8	71.34	1056.10	< 0.0001	significant
A	*72.91*	*1*	*72.91*	*1079.28*	*< 0.0001*	
B	*126.46*	*1*	*126.46*	*1872.01*	*< 0.0001*	
C	*103.46*	*1*	*103.46*	*1531.59*	*< 0.0001*	
D	*30.66*	*1*	*30.66*	*453.90*	*< 0.0001*	
AB	*29.93*	*1*	*29.93*	*443.01*	*< 0.0001*	
AC	*128.50*	*1*	*128.50*	*1902.15*	*< 0.0001*	
BC	*0.074*	*1*	*0.074*	*1.09*	*0.3075*	
ABC	*78.75*	*1*	*78.75*	*1165.76*	*< 0.0001*	
Residual	1.49	22	0.068			
Cor Total	572.25	31				

The Model F-value of 1056.10 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

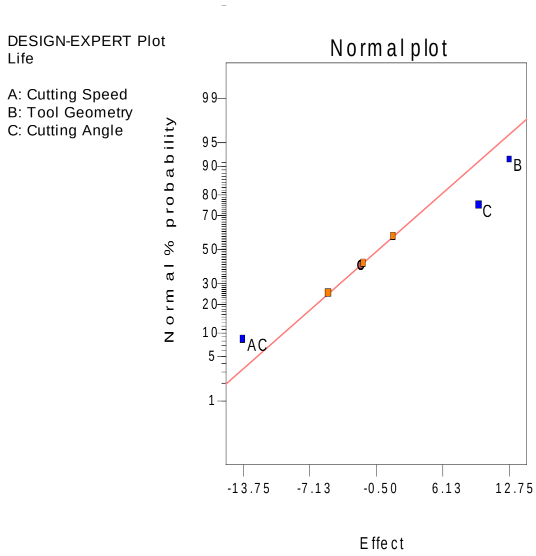
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case A, B, C, D, AB, AC, ABC are significant model terms.

Blocking does not change the results of Problem 6-15.

7.4. Consider the data from the first replicate of Problem 6.1. Suppose that these observations could not all be run using the same bar stock. Set up a design to run these observations in two blocks of four observations each with *ABC* confounded. Analyze the data.

Block 1	Block 2
(1)	*a*
ab	*b*
ac	*c*
bc	*abc*

From the normal probability plot of effects, *B*, *C*, and the *AC* interaction are significant. Factor *A* was included in the analysis of variance to preserve 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**	
Block	91.13	1	91.13			
Model	896.50	4	224.13	7.32	0.1238	not significant
A	*3.13*	*1*	*3.13*	*0.10*	*0.7797*	
B	*325.12*	*1*	*325.12*	*10.62*	*0.0827*	
C	*190.12*	*1*	*190.12*	*6.21*	*0.1303*	
AC	*378.13*	*1*	*378.13*	*12.35*	*0.0723*	
Residual	61.25	2	30.62			
Cor Total	1048.88	7				

The "Model F-value" of 7.32 implies the model is not significant relative to the noise. There is a 12.38 % chance that a "Model F-value" this large could occur due to noise.

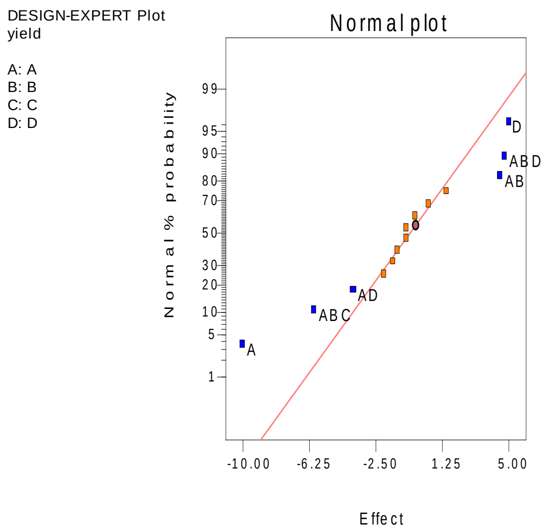
Values of "Prob > F" less than 0.0500 indicate model terms are significant.
In this case there are no significant model terms.

This design identifies the same significant factors as Problem 6.1.

7.5. Consider the data from the first replicate of Problem 6.7. Construct a design with two blocks of eight observations each with *ABCD* confounded. Analyze the data.

Block 1	Block 2
(1)	*a*
ab	*b*
ac	*c*
bc	*d*
ad	*abc*
bd	*abd*
cd	*acd*
abcd	*bcd*

The significant effects are identified in the normal probability plot of effects below:



AC, *BC*, and *BD* were included in the model to preserve hierarchy.

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**	
Block	42.25	1	42.25			
Model	892.25	11	81.11	9.64	0.0438	significant
A	*400.00*	*1*	*400.00*	*47.52*	*0.0063*	
B	*2.25*	*1*	*2.25*	*0.27*	*0.6408*	
C	*2.25*	*1*	*2.25*	*0.27*	*0.6408*	
D	*100.00*	*1*	*100.00*	*11.88*	*0.0410*	
AB	*81.00*	*1*	*81.00*	*9.62*	*0.0532*	
AC	*1.00*	*1*	*1.00*	*0.12*	*0.7531*	
AD	*56.25*	*1*	*56.25*	*6.68*	*0.0814*	
BC	*6.25*	*1*	*6.25*	*0.74*	*0.4522*	
BD	*9.00*	*1*	*9.00*	*1.07*	*0.3772*	
ABC	*144.00*	*1*	*144.00*	*17.11*	*0.0256*	
ABD	*90.25*	*1*	*90.25*	*10.72*	*0.0466*	
Residual	25.25	3	8.42			
Cor Total	959.75	15				

The Model F-value of 9.64 implies the model is significant. There is only a 4.38% chance that a "Model F-Value" this large could occur due to noise.

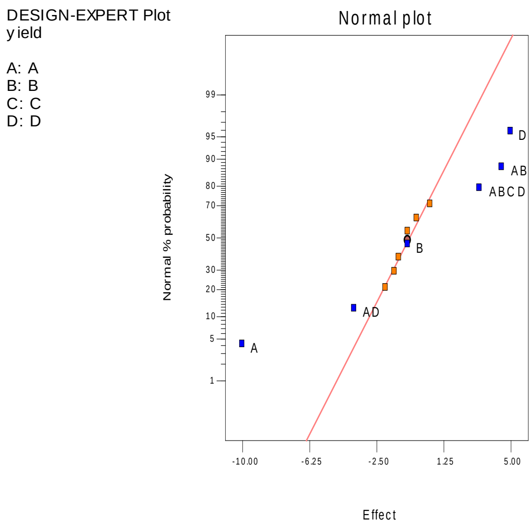
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, D, ABC, ABD are significant model terms.

7.6. Repeat Problem 7.5 assuming that four blocks are required. Confound *ABD* and *ABC* (and consequently *CD*) with blocks.

The block assignments are shown in the table below. The normal probability plot of effects identifies factors *A* and *D*, and the interactions *AB*, *AD*, and the *ABCD* as strong candidates for the model. For hierarchal purposes, factor *B* was included in the model; however, hierarchy is not preserved for the *ABCD* interaction allowing an estimate for error.

Block 1	Block 2	Block 3	Block 4

(1)	*ac*	*c*	*a*
ab	*bc*	*abc*	*b*
acd	*d*	*ad*	*cd*
bcd	*abd*	*bd*	*abcd*



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**	
Block	243.25	3	81.08			
Model	681.75	6	113.63	19.62	0.0011	significant
A	*400.00*	*1*	*400.00*	*69.06*	*0.0002*	
B	*2.25*	*1*	*2.25*	*0.39*	*0.5560*	
D	*100.00*	*1*	*100.00*	*17.27*	*0.0060*	
AB	*81.00*	*1*	*81.00*	*13.99*	*0.0096*	
AD	*56.25*	*1*	*56.25*	*9.71*	*0.0207*	
ABCD	*42.25*	*1*	*42.25*	*7.29*	*0.0355*	
Residual	34.75	6	5.79			
Cor Total	959.75	15				

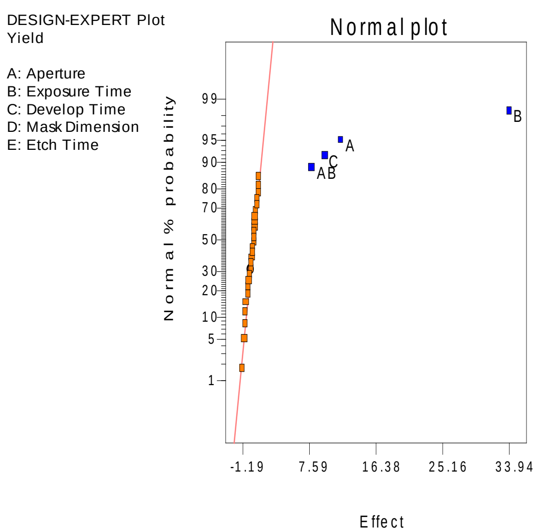
The Model F-value of 19.62 implies the model is significant. There is only a 0.11% chance that a "Model F-Value" this large could occur due to noise.

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

7.7. Using the data from the 2^5 design in Problem 6.26, construct and analyze a design in two blocks with *ABCDE* confounded with blocks.

Block 1	Block 1	Block 2	Block 2
(1)	*ae*	*a*	*e*
ab	*be*	*b*	*abe*
ac	*ce*	*c*	*ace*
bc	*abce*	*abc*	*bce*
ad	*de*	*d*	*ade*
bd	*abde*	*abd*	*bde*
cd	*acde*	*acd*	*cde*
abcd	*bcde*	*bcd*	*abcde*

The normal probability plot of effects identifies factors *A*, *B*, *C*, and the *AB* interaction as being significant. This is confirmed with the analysis of variance.



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**	
Block	0.28	1	0.28			
Model	11585.13	4	2896.28	958.51	< 0.0001	significant
A	*1116.28*	*1*	*1116.28*	*369.43*	*< 0.0001*	
B	*9214.03*	*1*	*9214.03*	*3049.35*	*< 0.0001*	
C	*750.78*	*1*	*750.78*	*248.47*	*< 0.0001*	
AB	*504.03*	*1*	*504.03*	*166.81*	*< 0.0001*	
Residual	78.56	26	3.02			
Cor Total	11663.97	31				

The Model F-value of 958.51 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

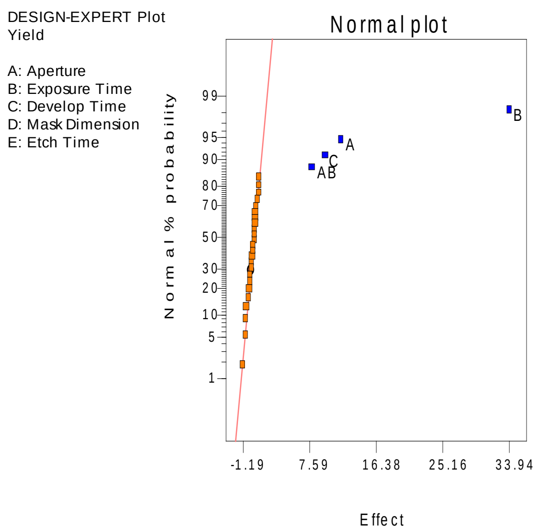
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, C, AB are significant model terms.

7.8. Repeat Problem 7.7 assuming that four blocks are necessary. Suggest a reasonable confounding scheme.

Use *ABC* and *CDE*, and consequently *ABDE*. The four blocks follow.

Block 1	Block 2	Block 3	Block 4
(1)	*a*	*ac*	*c*
ab	*b*	*bc*	*abc*
acd	*cd*	*d*	*ad*
bcd	*abcd*	*abd*	*bd*
ace	*ce*	*e*	*ae*
bce	*abce*	*abe*	*be*
de	*ade*	*acde*	*cde*
abde	*bde*	*bcde*	*abcde*

The normal probability plot of effects identifies the same significant effects as in Problem 7.7.



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**	
Block	13.84	3	4.61			
Model	11585.13	4	2896.28	1069.40	< 0.0001	significant
A	*1116.28*	*1*	*1116.28*	*412.17*	*< 0.0001*	
B	*9214.03*	*1*	*9214.03*	*3402.10*	*< 0.0001*	
C	*750.78*	*1*	*750.78*	*277.21*	*< 0.0001*	
AB	*504.03*	*1*	*504.03*	*186.10*	*< 0.0001*	
Residual	65.00	24	2.71			
Cor Total	11663.97	31				

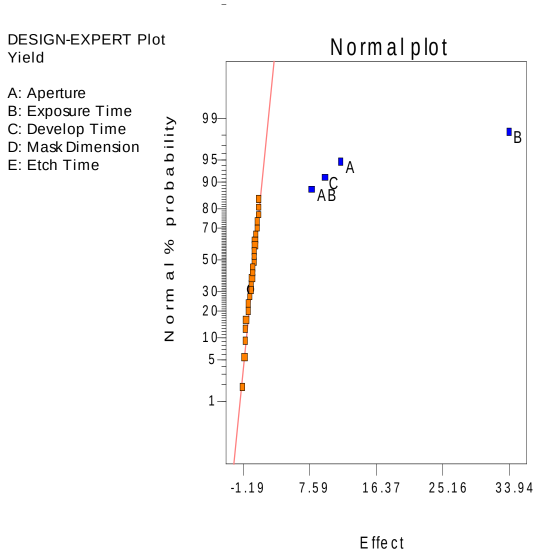
The Model F-value of 1069.40 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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

7.9. Consider the data from the 2^5 design in Problem 6.26. Suppose that it was necessary to run this design in four blocks with *ACDE* and *BCD* (and consequently *ABE*) confounded. Analyze the data from this design.

Block 1	Block 2	Block 3	Block 4
(1)	*a*	*b*	*c*
ae	*e*	*abe*	*ace*
cd	*acd*	*bcd*	*d*
abc	*bc*	*ac*	*ab*
acde	*cde*	*abcde*	*ade*
bce	*abce*	*ce*	*be*
abd	*bd*	*ad*	*abcd*
bde	*abde*	*de*	*bcde*

Even with four blocks, the same effects are identified as significant per the normal probability plot and analysis of variance below:



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**	
Block	2.59	3	0.86			
Model	11585.13	4	2896.28	911.62	< 0.0001	significant
A	*1116.28*	*1*	*1116.28*	*351.35*	*< 0.0001*	
B	*9214.03*	*1*	*9214.03*	*2900.15*	*< 0.0001*	
C	*750.78*	*1*	*750.78*	*236.31*	*< 0.0001*	
AB	*504.03*	*1*	*504.03*	*158.65*	*< 0.0001*	
Residual	76.25	24	3.18			
Cor Total	11663.97	31				

The Model F-value of 911.62 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.

7.10. Consider the fill height deviation experiment in Problem 6.20. Suppose that each replicate was run on a separate day. Analyze the data assuming that the days are blocks.

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	
Block	1.00	1	1.00			
Model	70.75	4	17.69	28.30	< 0.0001	significant
A	*36.00*	*1*	*36.00*	*57.60*	*< 0.0001*	
B	*20.25*	*1*	*20.25*	*32.40*	*0.0002*	
C	*12.25*	*1*	*12.25*	*19.60*	*0.0013*	
AB	*2.25*	*1*	*2.25*	*3.60*	*0.0870*	
Residual	6.25	10	0.62			
Cor Total	78.00	15				

The Model F-value of 28.30 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.

The analysis is very similar to the original analysis in chapter 6. The same effects are significant.

7.11. Consider the fill height deviation experiment in Problem 6.20. Suppose that only four runs could be made on each shift. Set up a design with *ABC* confounded in replicate 1 and *AC* confounded in replicate 2. Analyze the data and comment on your findings.

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	
Block	1.50	3	0.50			
Model	70.75	4	17.69	24.61	0.0001	significant
A	*36.00*	*1*	*36.00*	*50.09*	*0.0001*	
B	*20.25*	*1*	*20.25*	*28.17*	*0.0007*	
C	*12.25*	*1*	*12.25*	*17.04*	*0.0033*	
AB	*2.25*	*1*	*2.25*	*3.13*	*0.1148*	
Residual	5.75	8	0.72			
Cor Total	78.00	15				

The Model F-value of 24.61 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.

The analysis is very similar to the original analysis of Problem 6.20 and that of problem 7.10. The *AB* interaction is less significant in this scenario.

7.12. Consider the putting experiment in Problem 6.21. Analyze the data considering each replicate as a block.

The analysis is similar to that of Problem 6.21. Blocking has not changed the significant factors, however, the residual plots show that the normality assumption has been violated. The transformed data also has similar analysis to the transformed data of Problem 6.21. The ANOVA shown is for the transformed data.

Design Expert Output

Response: Distance from cupTransform:Square root Constant: 0
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	13.50	6	2.25			
Model	37.26	2	18.63	7.83	0.0007	significant
A	21.61	1	21.61	9.08	0.0033	
B	15.64	1	15.64	6.57	0.0118	
Residual	245.13	103	2.38			
Cor Total	295.89	111				

The Model F-value of 7.83 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.

7.13. Using the data from the 2^4 design in Problem 6.22, construct and analyze a design in two blocks with ABCD confounded with blocks.

Design Expert Output

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

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	2.500E-005	1	2.500E-005			
Model	0.24	4	0.059	32.33	< 0.0001	significant
A	*0.10*	*1*	*0.10*	*56.26*	*< 0.0001*	
C	*0.070*	*1*	*0.070*	*38.59*	*< 0.0001*	
D	*0.051*	*1*	*0.051*	*27.82*	*0.0004*	
AC	*0.012*	*1*	*0.012*	*6.65*	*0.0275*	
Residual	0.018	10	1.820E-003			
Cor Total	0.25	15				

The Model F-value of 32.33 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.

The analysis is similar to that of Problem 6.22. The significant effects are *A*, *C*, *D* and *AC*.

7.14. Consider the direct mail experiment in Problem 6.24. Suppose that each group of customers is in different parts of the country. Support an appropriate analysis for the experiment.

Set up each Group (replicate) as a geographic region. The analysis is similar to that of Problem 6.24. Factors *A* and *B* are included to achieve a hierarchical model.

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**	
Block	0.25	1	0.25			
Model	241.75	6	40.29	11.62	0.0014	significant
A	*12.25*	*1*	*12.25*	*3.53*	*0.0970*	
B	*2.25*	*1*	*2.25*	*0.65*	*0.4439*	
C	*36.00*	*1*	*36.00*	*10.38*	*0.0122*	
AB	*42.25*	*1*	*42.25*	*12.18*	*0.0082*	
AC	*100.00*	*1*	*100.00*	*28.83*	*0.0007*	
BC	*49.00*	*1*	*49.00*	*14.13*	*0.0056*	
Residual	27.75	8	3.47			
Cor Total	269.75	15				

The Model F-value of 11.62 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 C, AB, AC, BC are significant model terms.

7.15. Consider the isatin yield experiment in Problem 6.38. Set up the 2^4 experiment in this problem in two blocks with ABCD confounded. Analyze the data from this design. Is the block effect large?

The block effect is very small.

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**	
Block	1.406E-003	1	1.406E-003			
Model	0.55	3	0.18	4.15	0.0340	significant
B-Reaction time	*1.806E-003*	*1*	*1.806E-003*	*0.041*	*0.8440*	
D-Reaction temperature	*0.30*	*1*	*0.30*	*6.74*	*0.0249*	
BD	*0.25*	*1*	*0.25*	*5.68*	*0.0363*	
Residual	0.49	11	0.044			
Cor Total	1.04	15				

The Model F-value of 4.15 implies the model is significant. There is only a 3.40% 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, BD are significant model terms.

7.16. The experiment in Problem 6.39 is a 2^5 factorial. Suppose that this design had been run in four blocks of eight runs each.

(a) Recommend a blocking scheme and set up the design.

Interactions *ABC* and *BDE* are confounded with the blocks such that:

Block	*ABC*	*BDE*
1	-	+
2	+	-
3	-	-

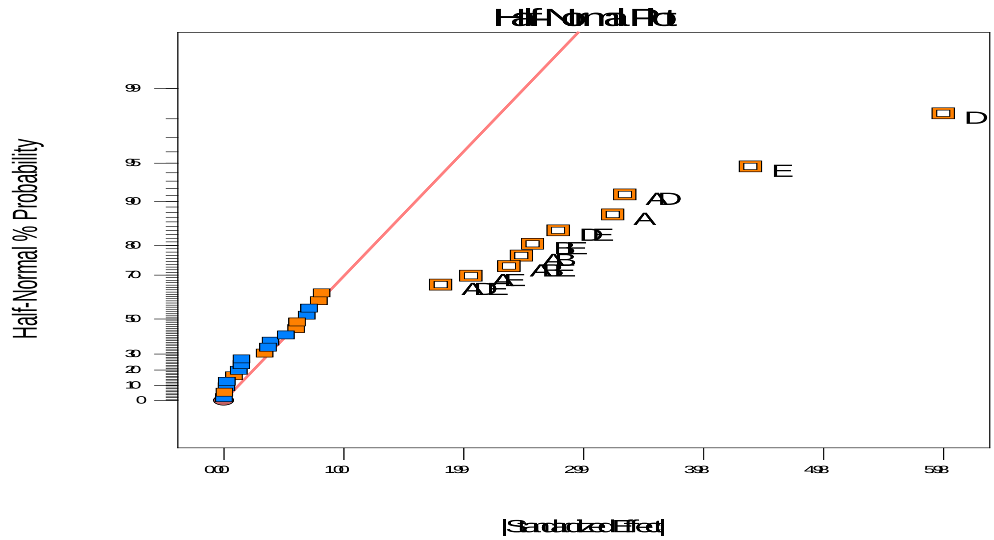
4	+	+

Note, the *ACDE* interaction is also confounded with the blocks. The experimental runs with the blocks are shown below.

Block	A	B	C	D	E	y
Block 1	-1	-1	-1	-1	-1	8.11
Block 2	1	-1	-1	-1	-1	5.56
Block 4	-1	1	-1	-1	-1	5.77
Block 3	1	1	-1	-1	-1	5.82
Block 2	-1	-1	1	-1	-1	9.17
Block 1	1	-1	1	-1	-1	7.8
Block 3	-1	1	1	-1	-1	3.23
Block 4	1	1	1	-1	-1	5.69
Block 3	-1	-1	-1	1	-1	8.82
Block 4	1	-1	-1	1	-1	14.23
Block 2	-1	1	-1	1	-1	9.2
Block 1	1	1	-1	1	-1	8.94
Block 4	-1	-1	1	1	-1	8.68
Block 3	1	-1	1	1	-1	11.49
Block 1	-1	1	1	1	-1	6.25
Block 2	1	1	1	1	-1	9.12
Block 3	-1	-1	-1	-1	1	7.93
Block 4	1	-1	-1	-1	1	5
Block 2	-1	1	-1	-1	1	7.47
Block 1	1	1	-1	-1	1	12
Block 4	-1	-1	1	-1	1	9.86
Block 3	1	-1	1	-1	1	3.65
Block 1	-1	1	1	-1	1	6.4
Block 2	1	1	1	-1	1	11.61
Block 1	-1	-1	-1	1	1	12.43
Block 2	1	-1	-1	1	1	17.55
Block 4	-1	1	-1	1	1	8.87
Block 3	1	1	-1	1	1	25.38
Block 2	-1	-1	1	1	1	13.06
Block 1	1	-1	1	1	1	18.85
Block 3	-1	1	1	1	1	11.78
Block 4	1	1	1	1	1	26.05

(b) Analyze the data from this blocked design. Is blocking important?

Blocking does not appear to be important; however, if the *ADE* or *ABE* interaction had been chosen to define the blocks, then blocking would have appeared as important. The *ADE* and *ABE* are significant effects in the analysis below.



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**	
Block	2.58	3	0.86			
Model	879.62	11	79.97	45.38	< 0.0001	significant
A-A	*83.56*	*1*	*83.56*	*47.41*	*< 0.0001*	
B-B	*0.060*	*1*	*0.060*	*0.034*	*0.8553*	
D-D	*285.78*	*1*	*285.78*	*162.16*	*< 0.0001*	
E-E	*153.17*	*1*	*153.17*	*86.91*	*< 0.0001*	
AB	*48.93*	*1*	*48.93*	*27.76*	*< 0.0001*	
AD	*88.88*	*1*	*88.88*	*50.43*	*< 0.0001*	
AE	*33.76*	*1*	*33.76*	*19.16*	*0.0004*	
BE	*52.71*	*1*	*52.71*	*29.91*	*< 0.0001*	
DE	*61.80*	*1*	*61.80*	*35.07*	*< 0.0001*	
ABE	*44.96*	*1*	*44.96*	*25.51*	*< 0.0001*	
ADE	*26.01*	*1*	*26.01*	*14.76*	*0.0013*	
Residual	29.96	17	1.76			
Cor Total	912.16	31				

The Model F-value of 45.38 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.

7.17. Repeat Problem 7.16 using a design in two blocks.

(a) Recommend a blocking scheme and set up the design.

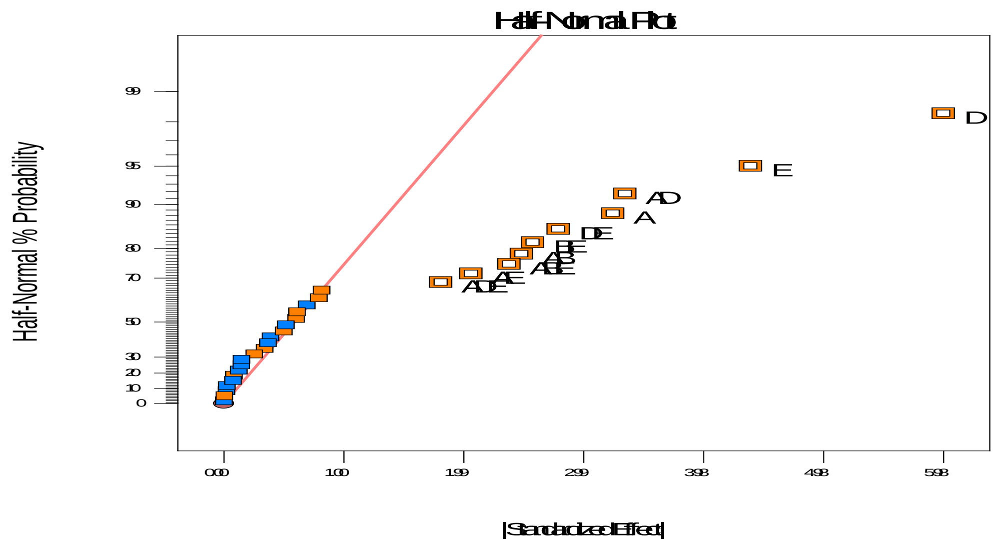
Interaction *ABCDE* is confounded with the blocks. The design is shown below.

Block	A	B	C	D	E	y
Block 1	-1	-1	-1	-1	-1	8.11
Block 2	1	-1	-1	-1	-1	5.56
Block 2	-1	1	-1	-1	-1	5.77

Block 1	1	1	-1	-1	-1	5.82
Block 2	-1	-1	1	-1	-1	9.17
Block 1	1	-1	1	-1	-1	7.8
Block 1	-1	1	1	-1	-1	3.23
Block 2	1	1	1	-1	-1	5.69
Block 2	-1	-1	-1	1	-1	8.82
Block 1	1	-1	-1	1	-1	14.23
Block 1	-1	1	-1	1	-1	9.2
Block 2	1	1	-1	1	-1	8.94
Block 1	-1	-1	1	1	-1	8.68
Block 2	1	-1	1	1	-1	11.49
Block 2	-1	1	1	1	-1	6.25
Block 1	1	1	1	1	-1	9.12
Block 2	-1	-1	-1	-1	1	7.93
Block 1	1	-1	-1	-1	1	5
Block 1	-1	1	-1	-1	1	7.47
Block 2	1	1	-1	-1	1	12
Block 1	-1	-1	1	-1	1	9.86
Block 2	1	-1	1	-1	1	3.65
Block 2	-1	1	1	-1	1	6.4
Block 1	1	1	1	-1	1	11.61
Block 1	-1	-1	-1	1	1	12.43
Block 2	1	-1	-1	1	1	17.55
Block 2	-1	1	-1	1	1	8.87
Block 1	1	1	-1	1	1	25.38
Block 2	-1	-1	1	1	1	13.06
Block 1	1	-1	1	1	1	18.85
Block 1	-1	1	1	1	1	11.78
Block 2	1	1	1	1	1	26.05

(b) Analyze the data from this blocked design. Is blocking important?

The analysis below shows that the blocking does not appear to be very important.



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**	
Block	4.04	1	4.04			
Model	879.62	11	79.97	53.31	< 0.0001	significant
A-A	*83.56*	*1*	*83.56*	*55.71*	*< 0.0001*	
B-B	*0.060*	*1*	*0.060*	*0.040*	*0.8431*	
D-D	*285.78*	*1*	*285.78*	*190.54*	*< 0.0001*	
E-E	*153.17*	*1*	*153.17*	*102.12*	*< 0.0001*	
AB	*48.93*	*1*	*48.93*	*32.62*	*< 0.0001*	
AD	*88.88*	*1*	*88.88*	*59.26*	*< 0.0001*	
AE	*33.76*	*1*	*33.76*	*22.51*	*0.0001*	
BE	*52.71*	*1*	*52.71*	*35.14*	*< 0.0001*	
DE	*61.80*	*1*	*61.80*	*41.20*	*< 0.0001*	
ABE	*44.96*	*1*	*44.96*	*29.98*	*< 0.0001*	
ADE	*26.01*	*1*	*26.01*	*17.34*	*0.0005*	
Residual	28.50	19	1.50			
Cor Total	912.16	31				

The Model F-value of 53.31 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.

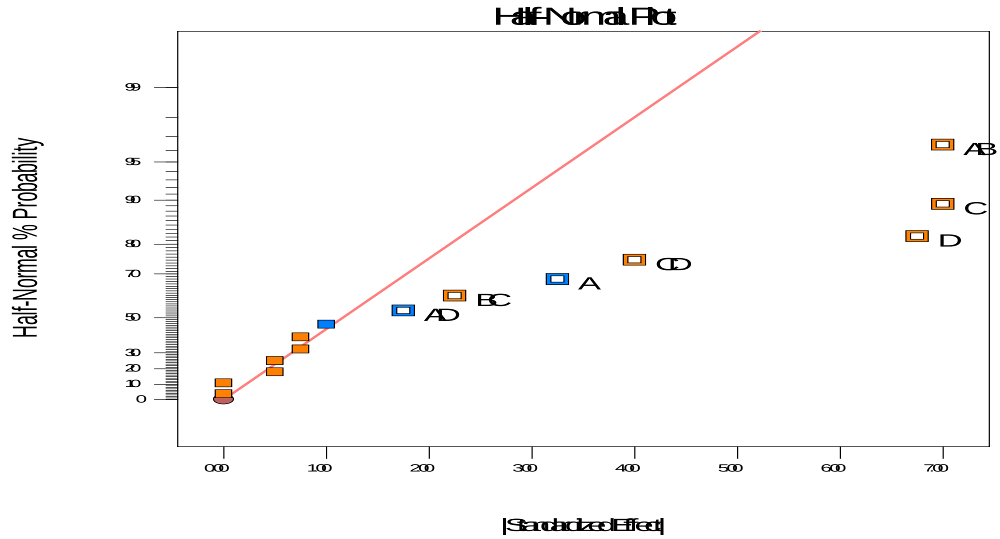
7.18. The design in Problem 6.40 is a 2^4 factorial. Set up this experiment in two blocks with ABCD confounded. Analyze the data from this design. Is the block effect large?

The runs for the experiment are shown below with the corresponding blocks.

Run	Block	Glucose (g dm^{-3})	NH_4NO_3 (g dm^{-3})	$FeSO_4$ (g dm^{-3} x 10^{-4})	$MnSO_4$ (g dm^{-3} x 10^{-2})	y (CMC)$^{-1}$
1	Block 2	20.00	2.00	6.00	4.00	23
2	Block 1	60.00	2.00	6.00	4.00	15
3	Block 1	20.00	6.00	6.00	4.00	16
4	Block 2	60.00	6.00	6.00	4.00	18

5	Block 1	20.00	2.00	30.00	4.00	25
6	Block 2	60.00	2.00	30.00	4.00	16
7	Block 2	20.00	6.00	30.00	4.00	17
8	Block 1	60.00	6.00	30.00	4.00	26
9	Block 1	20.00	2.00	6.00	20.00	28
10	Block 2	60.00	2.00	6.00	20.00	16
11	Block 2	20.00	6.00	6.00	20.00	18
12	Block 1	60.00	6.00	6.00	20.00	21
13	Block 2	20.00	2.00	30.00	20.00	36
14	Block 1	60.00	2.00	30.00	20.00	24
15	Block 1	20.00	6.00	30.00	20.00	33
16	Block 2	60.00	6.00	30.00	20.00	34

The analysis of the experiment shown below identifies the contribution of the blocks. By reducing the SS_E and MS_E, the *AD* and *CD* interactions now appear to be significant.



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**	
Block	6.25	1	6.25			
Model	713.00	8	89.13	50.93	< 0.0001	significant
A-Glucose	*42.25*	*1*	*42.25*	*24.14*	*0.0027*	
B-NH4NO3	*0.000*	*1*	*0.000*	*0.000*	*1.0000*	
C-FeSO4	*196.00*	*1*	*196.00*	*112.00*	*< 0.0001*	
D-MnSO4	*182.25*	*1*	*182.25*	*104.14*	*< 0.0001*	
AB	*196.00*	*1*	*196.00*	*112.00*	*< 0.0001*	
AD	*12.25*	*1*	*12.25*	*7.00*	*0.0382*	
BC	*20.25*	*1*	*20.25*	*11.57*	*0.0145*	
CD	*64.00*	*1*	*64.00*	*36.57*	*0.0009*	
Residual	10.50	6	1.75			
Cor Total	729.75	15				

The Model F-value of 50.93 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, BC, CD are significant model terms.

7.19. The design in Problem 6.42 is a 2^3 factorial replicated twice. Suppose that each replicate was a block. Analyze all of the responses from this blocked design. Are the results comparable to those from Problem 6.42? Is the block effect large?

The block effect is not large and does not appear to be important for the analysis on any of the four the responses as shown below. The results are comparable to those from Problem 6.42.

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	
Block	0.000	1	0.000			
Model	12.19	7	1.74	2300.73	< 0.0001	significant
A-Apatite	*10.99*	*1*	*10.99*	*14514.07*	*< 0.0001*	
B-pH	*0.35*	*1*	*0.35*	*459.75*	*< 0.0001*	
C-Pb	*0.27*	*1*	*0.27*	*350.30*	*< 0.0001*	
AB	*0.36*	*1*	*0.36*	*475.47*	*< 0.0001*	
AC	*0.19*	*1*	*0.19*	*249.92*	*< 0.0001*	
BC	*0.022*	*1*	*0.022*	*29.72*	*0.0010*	
ABC	*0.020*	*1*	*0.020*	*25.89*	*0.0014*	
Residual	5.300E-003	7	7.571E-004			
Cor Total	12.20	15				

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	0.85	1	6.879E-003	0.84	0.87	
Block 1	0.000	1				
Block 2	0.000					
A-Apatite	-0.83	1	6.879E-003	-0.85	-0.81	1.00
B-pH	-0.15	1	6.879E-003	-0.16	-0.13	1.00
C-Pb	0.13	1	6.879E-003	0.11	0.15	1.00
AB	0.15	1	6.879E-003	0.13	0.17	1.00
AC	-0.11	1	6.879E-003	-0.13	-0.092	1.00
BC	0.037	1	6.879E-003	0.021	0.054	1.00
ABC	-0.035	1	6.879E-003	-0.051	-0.019	1.00

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	
Block	2.256E-003	1	2.256E-003			
Model	21.09	7	3.01	102.87	< 0.0001	significant
A-Apatite	*9.84*	*1*	*9.84*	*336.14*	*< 0.0001*	
B-pH	*8.14*	*1*	*8.14*	*277.85*	*< 0.0001*	
C-Pb	*1.12*	*1*	*1.12*	*38.19*	*0.0005*	
AB	*0.61*	*1*	*0.61*	*20.91*	*0.0026*	
AC	*1.17*	*1*	*1.17*	*40.01*	*0.0004*	
BC	*0.098*	*1*	*0.098*	*3.33*	*0.1106*	
ABC	*0.11*	*1*	*0.11*	*3.66*	*0.0972*	
Residual	0.20	7	0.029			
Cor Total	21.30	15				

The Model F-value of 102.87 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, AC are significant model terms.

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	5.05	1	0.043	4.95	5.15	
Block 1	-0.012	1				
Block 2	0.012					
A-Apatite	0.78	1	0.043	0.68	0.89	1.00
B-pH	-0.71	1	0.043	-0.81	-0.61	1.00
C-Pb	-0.26	1	0.043	-0.37	-0.16	1.00
AB	0.20	1	0.043	0.094	0.30	1.00
AC	-0.27	1	0.043	-0.37	-0.17	1.00
BC	-0.078	1	0.043	-0.18	0.023	1.00
ABC	-0.082	1	0.043	-0.18	0.019	1.00

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	
Block	2.250E-004	1	2.250E-004			
Model	4.01	7	0.57	937.82	< 0.0001	significant
A-Apatite	*2.45*	*1*	*2.45*	*4010.43*	*< 0.0001*	
B-pH	*0.27*	*1*	*0.27*	*434.29*	*< 0.0001*	
C-Pb	*0.54*	*1*	*0.54*	*884.58*	*< 0.0001*	
AB	*0.17*	*1*	*0.17*	*275.25*	*< 0.0001*	
AC	*0.50*	*1*	*0.50*	*825.43*	*< 0.0001*	
BC	*0.036*	*1*	*0.036*	*59.11*	*0.0001*	
ABC	*0.046*	*1*	*0.046*	*75.69*	*< 0.0001*	
Residual	4.275E-003	7	6.107E-004			
Cor Total	4.01	15				

The Model F-value of 937.82 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, AC, BC, ABC are significant model terms.

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	0.42	1	6.178E-003	0.40	0.43	
Block 1	3.750E-003	1				
Block 2	-3.750E-003					
A-Apatite	-0.39	1	6.178E-003	-0.41	-0.38	1.00
B-pH	0.13	1	6.178E-003	0.11	0.14	1.00
C-Pb	0.18	1	6.178E-003	0.17	0.20	1.00
AB	-0.10	1	6.178E-003	-0.12	-0.088	1.00
AC	-0.18	1	6.178E-003	-0.19	-0.16	1.00
BC	-0.048	1	6.178E-003	-0.062	-0.033	1.00
ABC	0.054	1	6.178E-003	0.039	0.068	1.00

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	
Block	2.025E-003	1	2.025E-003			
Model	20.44	7	2.92	1494.46	< 0.0001	significant
A-Apatite	*8.15*	*1*	*8.15*	*4172.37*	*< 0.0001*	
B-pH	*8.82*	*1*	*8.82*	*4515.27*	*< 0.0001*	
C-Pb	*0.084*	*1*	*0.084*	*43.05*	*0.0003*	
AB	*3.24*	*1*	*3.24*	*1658.50*	*< 0.0001*	
AC	*0.014*	*1*	*0.014*	*7.37*	*0.0300*	

BC	*0.13*	*1*	*0.13*	*64.51*	*< 0.0001*
ABC	*2.250E-004*	*1*	*2.250E-004*	*0.12*	*0.7443*
Residual	0.014	7	1.954E-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	
Block 1	0.011	1				
Block 2	-0.011					
A-Apatite	0.71	1	0.011	0.69	0.74	1.00
B-pH	-0.74	1	0.011	-0.77	-0.72	1.00
C-Pb	-0.073	1	0.011	-0.099	-0.046	1.00
AB	-0.45	1	0.011	-0.48	-0.42	1.00
AC	-0.030	1	0.011	-0.056	-3.871E-003	1.00
BC	0.089	1	0.011	0.063	0.11	1.00
ABC	3.750E-003	1	0.011	-0.022	0.030	1.00

7.20. Design an experiment for confounding a 2^6 factorial in four blocks. Suggest an appropriate confounding scheme, different from the one shown in Table 7.8.

We choose *ABCE* and *ABDF*, which also confounds *CDEF*.

Block 1	Block 2	Block 3	Block 4
a	*c*	*ac*	(1)
b	*abc*	*bc*	*ab*
cd	*ad*	*d*	*acd*
abcd	*bd*	*abd*	*bcd*
ace	*e*	*ae*	*ce*
bce	*abe*	*be*	*abce*
de	*acde*	*cde*	*ade*
abde	*bcde*	*abcde*	*bde*
cf	*af*	*f*	*acf*
abcf	*bf*	*abf*	*bcf*
adf	*cdf*	*acdf*	*df*
bdf	*abcdf*	*bcdf*	*abdf*
ef	*acef*	*cef*	*aef*
abef	*bcef*	*abcef*	*bef*
acdef	*def*	*adef*	*cdef*
bcdef	*abdef*	*bdef*	*abcdef*

7.21. Consider the 2^6 design in eight blocks of eight runs each with *ABCD*, *ACE*, and *ABEF* as the independent effects chosen to be confounded with blocks. Generate the design. Find the other effects confound with blocks.

Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8
b	*abc*	*a*	*c*	*ac*	(1)	*bc*	*ab*
acd	*d*	*bcd*	*abd*	*bd*	*abcd*	*ad*	*cd*
ce	*ae*	*abce*	*be*	*abe*	*bce*	*e*	*ace*
abde	*bcde*	*de*	*acde*	*cde*	*ade*	*abcde*	*bde*
abcf	*bf*	*cf*	*af*	*f*	*acf*	*abf*	*bcf*
df	*acdf*	*abdf*	*bcdf*	*abcdf*	*bdf*	*cdf*	*adf*
aef	*cef*	*bef*	*abcef*	*bcef*	*abef*	*acef*	*ef*
bcdef	*abdef*	*acdef*	*def*	*adef*	*cdef*	*bdef*	*abcdef*

The factors that are confounded with blocks are *ABCD, ABEF, ACE, BDE, CDEF, BCF*, and *ADF*.

7.22. Consider the 2^2 design in two blocks with *AB* confounded. Prove algebraically that $SS_{AB} = SS_{\text{Blocks}}$.

If *AB* is confounded, the two blocks are:

Block 1	Block 2
(1)	*a*
ab	*b*
(1) + *ab*	*a* + *b*

$$SS_{Blocks} = \frac{\left[(1)+ab\right]^2 + \left[a+b\right]^2}{2} - \frac{\left[(1)+ab+a+b\right]^2}{4}$$

$$SS_{Blocks} = \frac{(1)^2 + (ab)^2 + 2(1)ab + a^2 + b^2 + 2ab}{2}$$

$$-\frac{(1)^2 + (ab)^2 + a^2 + b^2 + 2(1)ab + 2(1)a + 2(1)b + 2a(ab) + 2b(ab) + 2ab}{4}$$

$$SS_{Blocks} = \frac{(1)^2 + (ab)^2 + a^2 + b^2 + 2(1)ab + 2ab - 2(1)a - 2(1)b - 2a(ab) - 2b(ab)}{4}$$

$$SS_{Blocks} = \frac{1}{4}\left[(1)+ab-a-b\right]^2 = SS_{AB}$$

7.23. Consider the data in Example 7.2. Suppose that all the observations in block 2 are increased by 20. Analyze the data that would result. Estimate the block effect. Can you explain its magnitude? Do blocks now appear to be an important factor? Are any other effect estimates impacted by the change you made in the data?

$$\text{Block Effect} = \bar{y}_{Block1} - \bar{y}_{Block2} = \frac{406}{8} - \frac{715}{8} = \frac{-309}{8} = -38.625$$

This is the block effect estimated in Example 7.2 plus the additional 20 units that were added to each observation in block 2. All other effects are the same.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	1870.56	1	1870.56	89.93
C	390.06	1	390.06	18.75
D	855.56	1	855.56	41.13
AC	1314.06	1	1314.06	63.18
AD	1105.56	1	1105.56	53.15
Blocks	5967.56	1	5967.56	
Error	187.56	9	20.8	
Total	11690.93	15		

Design Expert Output

Response: Filtration in gal/hr
ANOVA for Selected Factorial Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Block	5967.56	1	5967.56			
Model	5535.81	5	1107.16	53.13	< 0.0001	significant
A	*1870.56*	*1*	*1870.56*	*89.76*	*< 0.0001*	
C	*390.06*	*1*	*390.06*	*18.72*	*0.0019*	
D	*855.56*	*1*	*855.56*	*41.05*	*0.0001*	
AC	*1314.06*	*1*	*1314.06*	*63.05*	*< 0.0001*	
AD	*1105.56*	*1*	*1105.56*	*53.05*	*< 0.0001*	
Residual	187.56	9	20.84			
Cor Total	11690.94	15				

The Model F-value of 53.13 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.

7.24. Suppose that the data in Problem 6.1 we had confounded *ABC* in replicate I, *AB* in replicate II, and *BC* in replicate III. Construct the analysis of variance table.

	Replicate I (*ABC* Confounded)		Replicate II (*AB* Confounded)		Replicate III (*BC* Confounded)	
Block->	1	2	1	2	1	2
	(1)	*a*	(1)	*a*	(1)	*b*
	ab	*b*	*ab*	*b*	*bc*	*c*
	ac	*c*	*abc*	*ac*	*abc*	*ab*
	bc	*abc*	*c*	*bc*	*a*	*ac*

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	0.67	1	0.67	<1
B	770.67	1	770.67	20.77
C	280.17	1	280.17	7.55
AB (reps 1 and III)	25.00	1	25.00	<1
AC	468.17	1	468.17	12.62
BC (reps I and II)	22.56	1	22.56	<1
ABC (reps II and III)	0.06	1	0.06	<1
Blocks within replicates	119.25	3	39.75	
Replicates	0.58	2		
Error	408.21	11	37.11	
Total	2095.33	23		

7.25. Repeat the analysis of Problem 6.1 assuming that *ABC* was confounded with blocks in each replicate.

	Replicate I, II, and III (*ABC* Confounded)	
Block->	1	2
	(1)	*a*
	ab	*b*
	ac	*c*
	bc	*abc*

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	0.67	1	0.67	<1
B	770.67	1	770.67	22.38
C	280.17	1	280.17	8.14
AB	16.67	1	16.67	<1
AC	468.17	1	468.17	13.59
BC	48.17	1	48.17	1.40
Blocks (or *ABC*)	28.17	1	28.17	
Replicates/Lack of Fit	0.58	2		
Error	482.09	14	34.44	
Total	2095.33	23		

7.26. Suppose that in Problem 6.7 *ABCD* was confounded in replicate I and *ABC* was confounded in replicate II. Perform the statistical analysis of variance.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
A	657.03	1	657.03	84.89
B	13.78	1	13.78	1.78
C	57.78	1	57.78	7.46
D	124.03	1	124.03	16.02
AB	132.03	1	132.03	17.06
AC	3.78	1	3.78	<1
AD	38.28	1	38.28	4.95
BC	2.53	1	2.53	<1
BD	0.28	1	0.28	<1
CD	22.78	1	22.78	2.94
ABC	144.00	1	144.00	18.64
ABD	175.78	1	175.78	22.71
ACD	7.03	1	7.03	<1
BCD	7.03	1	7.03	<1
ABCD	10.56	1	10.56	1.36
Replicates	11.28	1	11.28	
Blocks	118.81	2	59.41	
Error	100.65	13	7.74	
Total	1627.47	31		

7.27. Construct a 2^3 design with *ABC* confounded in the first two replicates and *BC* confounded in the third. Outline the analysis of variance and comment on the information obtained.

	Replicate I (*ABC* Confounded)		Replicate II (*ABC* Confounded)		Replicate III (*BC* Confounded)	
Block->	1	2	1	2	1	2
	(1)	*a*	(1)	*a*	(1)	*b*
	ab	*b*	*ab*	*b*	*bc*	*c*
	ac	*c*	*ac*	*c*	*abc*	*ab*
	bc	*abc*	*bc*	*abc*	*a*	*ac*

Source of Variation	Degrees of Freedom
A	1
B	1
C	1
AB	1
AC	1
BC	1
ABC	1
Replicates	2
Blocks	3
Error	11
Total	23

This design provides “two-thirds” information on *BC* and “one-third” information on *ABC*.

Chapter 8
Two-Level Fractional Factorial Designs
Solutions

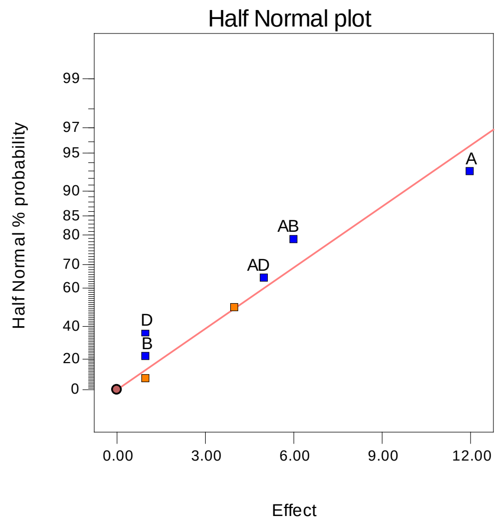
8.1. Suppose that in the chemical process development experiment in Problem 6.7, it was only possible to run a one-half fraction of the 2^4 design. Construct the design and perform the statistical analysis, using the data from replicate 1.

The required design is a 2^{4-1} with $I=ABCD$.

A	*B*	*C*	*D=ABC*		
-	-	-	-	(1)	90
+	-	-	+	*ad*	72
-	+	-	+	*bd*	87
+	+	-	-	*ab*	83
-	-	+	+	*cd*	99
+	-	+	-	*ac*	81
-	+	+	-	*bc*	88
+	+	+	+	*abcd*	80

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	-12	288	64.2857
Model	B	-1	2	0.446429
Model	C	4	32	7.14286
Model	D	-1	2	0.446429
Model	AB	6	72	16.0714
Model	AC	-1	2	0.446429
Model	AD	-5	50	11.1607
Error	BC	Aliased		
Error	BD	Aliased		
Error	CD	Aliased		
Error	ABC	Aliased		
Error	ABD	Aliased		
Error	ACD	Aliased		
Error	BCD	Aliased		
Error	ABCD	Aliased		
	Lenth's ME		22.5856	
	Lenth's SME 54.0516			



The largest effect is *A*. The next largest effects are the *AB* and *AD* interactions. A plausible tentative model would be *A*, *AB* and *AD*, along with *B* and *D* to preserve hierarchy.

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	414.00	5	82.80	4.87	0.1791	not significant
A	*288.00*	*1*	*288.00*	*16.94*	*0.0543*	
B	*2.00*	*1*	*2.00*	*0.12*	*0.7643*	
D	*2.00*	*1*	*2.00*	*0.12*	*0.7643*	
AB	*72.00*	*1*	*72.00*	*4.24*	*0.1758*	
AD	*50.00*	*1*	*50.00*	*2.94*	*0.2285*	
Residual	34.00	2	17.00			
Cor Total	448.00	7				

The "Model F-value" of 4.87 implies the model is not significant relative to the noise. There is a 17.91 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	4.12	R-Squared	0.9241
Mean	85.00	Adj R-Squared	0.7344
C.V.	4.85	Pred R-Squared	-0.2143
PRESS	544.00	Adeq Precision	6.441

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	85.00	1	1.46	78.73	91.27	
A-A	-6.00	1	1.46	-12.27	0.27	1.00
B-B	-0.50	1	1.46	-6.77	5.77	1.00
D-D	-0.50	1	1.46	-6.77	5.77	1.00
AB	3.00	1	1.46	-3.27	9.27	1.00
AD	-2.50	1	1.46	-8.77	3.77	1.00

Final Equation in Terms of Coded Factors:

yield	=
+85.00	
-6.00	* A
-0.50	* B
-0.50	* D
+3.00	* A * B
-2.50	* A * D

Final Equation in Terms of Actual Factors:

yield	=
+85.00000	
-6.00000	* A
-0.50000	* B
-0.50000	* D
+3.00000	* A * B
-2.50000	* A * D

The Design-Expert output indicates that we really only need the main effect of factor *A*. The updated analysis is shown below:

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	288.00	1	288.00	10.80	0.0167	significant
A	*288.00*	*1*	*288.00*	10.80	0.0167	
Residual	160.00	6	26.67			
Cor Total	448.00	7				

The Model F-value of 10.80 implies the model is significant. There is only a 1.67% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	5.16	R-Squared	0.6429
Mean	85.00	Adj R-Squared	0.5833
C.V.	6.08	Pred R-Squared	0.3651
PRESS	284.44	Adeq Precision	4.648

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	85.00	1	1.83	80.53	89.47	
A-A	-6.00	1	1.83	-10.47	-1.53	1.00

Final Equation in Terms of Coded Factors:

yield	=
+85.00	
-6.00	* A

Final Equation in Terms of Actual Factors:

yield	=
+85.00000	
-6.00000	* A

8.2. Suppose that in Problem 6.15, only a one-half fraction of the 2^4 design could be run. Construct the design and perform the analysis, using the data from replicate I.

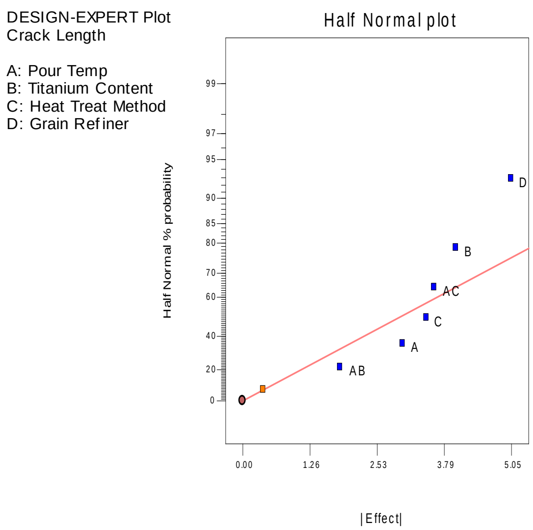
The required design is a 2^{4-1} with *I=ABCD*.

A	*B*	*C*	*D=ABC*		
-	-	-	-	(1)	1.71
+	-	-	+	*ad*	1.86
-	+	-	+	*bd*	1.79
+	+	-	-	*ab*	1.67
-	-	+	+	*cd*	1.81
+	-	+	-	*ac*	1.25
-	+	+	-	*bc*	1.46
+	+	+	+	*abcd*	0.85

Design Expert Output

Term	Effect	Effects	SumSqr	% Contribtn
Require	Intercept			
Model	A	3.0105	18.1262	11.4565
Model	B	4.011	32.1762	20.3366
Model	C	-3.4565	23.8948	15.1024
Model	D	5.051	51.0252	32.2499
Model	AB	1.8345	6.73078	4.25411
Model	AC	-3.603	25.9632	16.4097
Model	AD	0.3885	0.301865	0.19079
	Lenth's ME	19.5168		
	Lenth's SME	46.7074		

B, *D*, and *AC* + *BD* are the largest three effects. Now because the main effects of *B* and *D* are large, the large effect estimate for the *AC* + *BD* alias chain probably indicates that the *BD* interaction is important. It is also possible that the AB interaction is actually the CD interaction. This is not an easy decision. Additional experimental runs may be required to de-alias these two interactions.



Design Expert Output

Response: Crack Lengthin 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	157.92	6	26.32	87.19	0.0818	not significant
A	*18.13*	*1*	*18.13*	*60.05*	*0.0817*	
B	*32.18*	*1*	*32.18*	*106.59*	*0.0615*	
C	*23.89*	*1*	*23.89*	*79.16*	*0.0713*	
D	*51.03*	*1*	*51.03*	*169.03*	*0.0489*	
AB	*6.73*	*1*	*6.73*	*22.30*	*0.1329*	
BD	*25.96*	*1*	*25.96*	*86.01*	*0.0684*	
Residual	0.30	1	0.30			
Cor Total	158.22	7				

The Model F-value of 87.19 implies there is a 8.18% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.55	R-Squared	0.9981
Mean	12.04	Adj R-Squared	0.9866
C.V.	4.57	Pred R-Squared	0.8779
PRESS	19.32	Adeq Precision	25.110

	Coefficient		**Standard**	**95% CI**	**95% CI**

Factor	Estimate	DF	Error	Low	High	VIF
Intercept	12.04	1	0.19	9.57	14.50	
A-Pour Temp	1.51	1	0.19	-0.96	3.97	1.00
B-Titanium Content	2.01	1	0.19	-0.46	4.47	1.00
C-Heat Treat Method	-1.73	1	0.19	-4.20	0.74	1.00
D-Grain Refiner	2.53	1	0.19	0.057	4.99	1.00
AB	0.92	1	0.19	-1.55	3.39	1.00
BD	-1.80	1	0.19	-4.27	0.67	1.00

Final Equation in Terms of Coded Factors:

Crack Length =
+12.04
+1.51 * A
+2.01 * B
-1.73 * C
+2.53 * D
+0.92 * A * B
-1.80 * B * D

Final Equation in Terms of Actual Factors:

Crack Length =
+12.03500
+1.50525 * Pour Temp
+2.00550 * Titanium Content
-1.72825 * Heat Treat Method
+2.52550 * Grain Refiner
+0.91725 * Pour Temp * Titanium Content
-1.80150 * Titanium Content * Grain Refiner

8.3. Consider the plasma etch experiment described in Example 6.1. Suppose that only a one-half fraction of the design could be run. Set up the design and analyze the data.

Because Example 6.1 is a replicated 2^3 factorial experiment, a half fraction of this design is a 2^{3-1} with four runs. The experiment is replicates to assure an adequate estimate of the MS_E.

A	*B*	*C*=*AB*	Etch Rate (A/min)		Factor Low (-)	Levels High (+)
-	-	+	1037	*A* (Gap, cm)	0.80	1.20
-	-	+	1052	*B* (C_2F_6 flow, SCCM)	125	200
+	-	-	669	*C* (Power, W)	275	325
+	-	-	650			
-	+	-	633			
-	+	-	601			
+	+	+	729			
+	+	+	860			

The analysis shown below identifies all three main effects as significant. Because this is a resolution III design, the main effects are aliased with two factor interactions. The original analysis from Example 6-1 identifies factors *A*, *C*, and the *AC* interaction as significant. In our replicated half fraction experiment, factor *B* is aliased with the *AC* interaction. This problem points out the concerns of running small resolution III designs.

Design Expert Output

Response: Etch 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	2.225E+005	3	74169.79	31.61	0.0030	significant
A	*21528.13*	*1*	*21528.13*	*9.18*	*0.0388*	
B	*42778.13*	*1*	*42778.13*	*18.23*	*0.0130*	
C	*1.582E+005*	*1*	*1.582E+005*	*67.42*	*0.0012*	
Pure Error	9385.50	4	2346.37			
Cor Total	2.319E+005	7				

The Model F-value of 31.61 implies the model is significant. There is only a 0.30% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	48.44	R-Squared	0.9595
Mean	778.88	Adj R-Squared	0.9292
C.V.	6.22	Pred R-Squared	0.8381
PRESS	37542.00	Adeq Precision	12.481

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	778.88	1	17.13	731.33	826.42	
A-Gap	-51.88	1	17.13	-99.42	-4.33	1.00
B-C2F6 Flow	-73.13	1	17.13	-120.67	-25.58	1.00
C-Power	140.63	1	17.13	93.08	188.17	1.00

Final Equation in Terms of Coded Factors:

Etch Rate	=
+778.88	
-51.88	* A
-73.13	* B
+140.63	* C

Final Equation in Terms of Actual Factors:

Etch Rate	=
-332.37500	
-259.37500	* Gap
-1.95000	* C2F6 Flow
+5.62500	* Power

8.4. Problem 6.26 describes a process improvement study in the manufacturing process of an integrated circuit. Suppose that only eight runs could be made in this process. Set up an appropriate 2^{5-2} design and find the alias structure. Use the appropriate observations from Problem 6.26 as the observations in this design and estimate the factor effects. What conclusions can you draw?

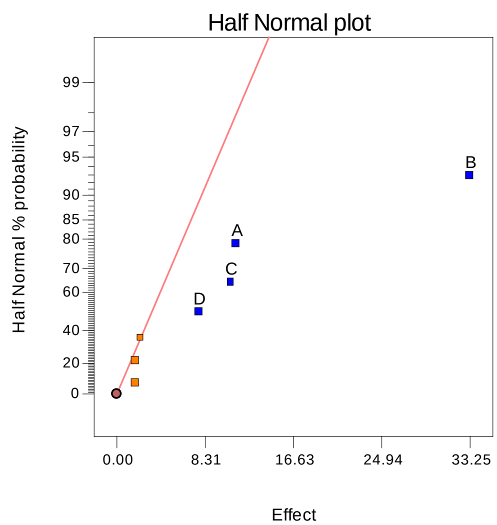
$$I = ABD = ACE = BCDE$$

A	*(ABD)*	*=BD*	*A*	*(ACE)*	*=CE*	*A*	*(BCDE)*	*=ABCDE*	*A=BD=CE=ABCDE*
B	*(ABD)*	*=AD*	*B*	*(ACE)*	*=ABCE*	*B*	*(BCDE)*	*=CDE*	*B=AD=ABCE=CDE*
C	*(ABD)*	*=ABCD*	*C*	*(ACE)*	*=AE*	*C*	*(BCDE)*	*=BDE*	*C=ABCD=AE=BDE*
D	*(ABD)*	*=AB*	*D*	*(ACE)*	*=ACDE*	*D*	*(BCDE)*	*=BCE*	*D=AB=ACDE=BCE*
E	*(ABD)*	*=ABDE*	*E*	*(ACE)*	*=AC*	*E*	*(BCDE)*	*=BCD*	*E=ABDE=AC=BCD*
BC	*(ABD)*	*=ACD*	*BC*	*(ACE)*	*=ABE*	*BC*	*(BCDE)*	*=DE*	*BC=ACD=ABE=DE*
BE	*(ABD)*	*=ADE*	*BE*	*(ACE)*	*=ABC*	*BE*	*(BCDE)*	*=CD*	*BE=ADE=ABC=CD*

A	*B*	*C*	*D=AB*	*E=AC*		
-	-	-	+	+	*de*	6
+	-	-	-	-	*a*	9
-	+	-	-	+	*be*	35
+	+	-	+	-	*abd*	50
-	-	+	+	-	*cd*	18
+	-	+	-	+	*ace*	22
-	+	+	-	-	*bc*	40
+	+	+	+	+	*abcde*	63

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	11.25	253.125	8.91953
Model	B	33.25	2211.13	77.9148
Model	C	10.75	231.125	8.1443
Model	D	7.75	120.125	4.23292
Error	E	2.25	10.125	0.356781
Error	BC	-1.75	6.125	0.215831
Error	BE	1.75	6.125	0.215831
	Lenth's ME	28.232		
	Lenth's SME	67.5646		



The main *A*, *B*, *C*, and *D* are large. However, recall that we are really estimating *A+BD+CE*, *B+AD*, *C+DE* and *D+AD*. There are other possible interpretations of the experiment because of the aliasing.

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	2815.50	4	703.88	94.37	0.0017	significant
A	*253.13*	*1*	*253.13*	*33.94*	*0.0101*	
B	*2211.12*	*1*	*2211.12*	*296.46*	*0.0004*	
C	*231.13*	*1*	*231.13*	*30.99*	*0.0114*	
D	*120.13*	*1*	*120.13*	*16.11*	*0.0278*	
Residual	22.38	3	7.46			
Cor Total	2837.88	7				

The Model F-value of 94.37 implies the model is significant. There is only a 0.17% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.73	R-Squared	0.9921
Mean	30.38	Adj R-Squared	0.9816
C.V.	8.99	Pred R-Squared	0.9439
PRESS	159.11	Adeq Precision	25.590

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	30.38	1	0.97	27.30	33.45	
A-Aperture	5.63	1	0.97	2.55	8.70	1.00
B-Exposure Time	16.63	1	0.97	13.55	19.70	1.00
C-Develop Time	5.37	1	0.97	2.30	8.45	1.00

D-Mask Dimension 3.87	1	0.97	0.80	6.95	1.00

Final Equation in Terms of Coded Factors:

```
        Yield   =
       +30.38
        +5.63    * A
       +16.63    * B
        +5.37    * C
        +3.87    * D
```

Final Equation in Terms of Actual Factors:

```
       Aperture   small
 Mask Dimension   Small
          Yield   =
       -6.00000
       +0.83125    * Exposure Time
       +0.71667    * Develop Time

       Aperture   large
 Mask Dimension   Small
          Yield   =
       +5.25000
       +0.83125    * Exposure Time
       +0.71667    * Develop Time

       Aperture   small
 Mask Dimension   Large
          Yield   =
       +1.75000
       +0.83125    * Exposure Time
       +0.71667    * Develop Time

       Aperture   large
 Mask Dimension   Large
          Yield   =
      +13.00000
       +0.83125    * Exposure Time
       +0.71667    * Develop Time
```

8.5. Continuation of Problem 8.4. Suppose you have made the eight runs in the 2^{5-2} design in Problem 8.4. What additional runs would be required to identify the factor effects that are of interest? What are the alias relationships in the combined design?

We could fold over the original design by changing the signs on the generators $D = AB$ and $E = AC$ to produce the following new experiment.

A	*B*	*C*	*D=-AB*	*E=-AC*		
-	-	-	-	-	(1)	7
+	-	-	+	+	*ade*	10
-	+	-	+	-	*bd*	32
+	+	-	-	+	*abe*	52
-	-	+	-	+	*ce*	15
+	-	+	+	-	*acd*	21
-	+	+	+	+	*bcde*	41
+	+	+	-	-	*abc*	60

A	(-ABD)	=-BD	A	(-ACE)	=-CE	A	(BCDE)	=ABCDE	A=-BD=-CE=ABCDE
B	(-ABD)	=-AD	B	(-ACE)	=-ABCE	B	(BCDE)	=CDE	B=-AD=-ABCE=CDE
C	(-ABD)	=-ABCD	C	(-ACE)	=-AE	C	(BCDE)	=BDE	C=-ABCD=-AE=BDE
D	(-ABD)	=-AB	D	(-ACE)	=-ACDE	D	(BCDE)	=BCE	D=-AB=-ACDE=BCE
E	(-ABD)	=-ABDE	E	(-ACE)	=-AC	E	(BCDE)	=BCD	E=-ABDE=-AC=BCD
BC	(-ABD)	=-ACD	BC	(-ACE)	=-ABE	BC	(BCDE)	=DE	BC=-ACD=-ABE=DE
BE	(-ABD)	=-ADE	BE	(-ACE)	=-ABC	BE	(BCDE)	=CD	BE=-ADE=-ABC=CD

Assuming all three factor and higher interactions to be negligible, all main effects can be separated from their two-factor interaction aliases in the combined design.

8.6. In the Example 6.6, a 2^4 factorial design was used to improve the response rate to a credit card marketing offer. Suppose that the researchers had used the 2^{4-1} fraction factorial design with $I=ABCD$ instead. Set up the design and select the responses for the runs from the full factorial data in Example 6.6. Analyze the data and draw conclusions. Compare your findings with those from the full factorial in Example 6.6.

Based on the Pseudo *p*-Value, the effects appear to be much less significant than those found in Example 6.6. The estimates for the Long-term Interest Rate, Account Opening Fee, and Annual Fee * Long-term Interest Rate are similar to the estimates found in Example 6.6; however, the other estimates are not as similar.

JMP Output
Response Response rate
Summary of Fit

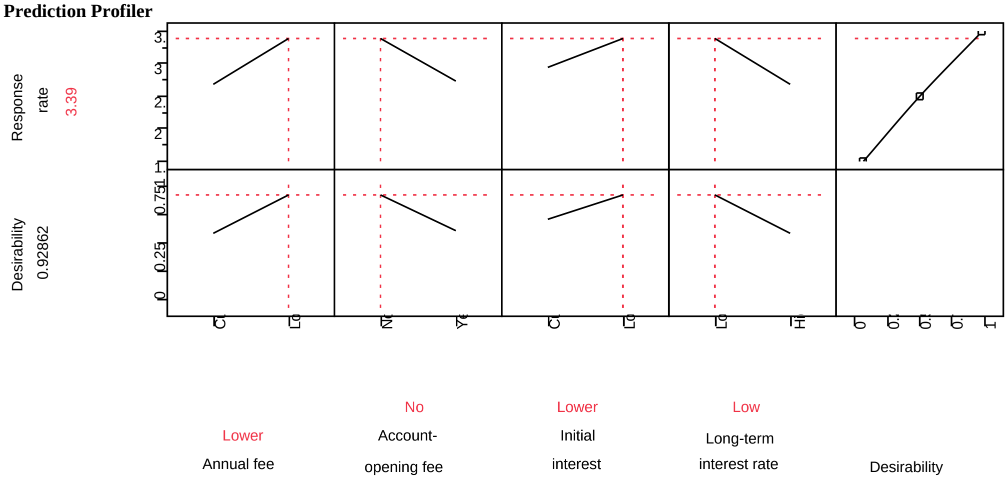
RSquare	1
RSquare Adj	.
Root Mean Square Error	.
Mean of Response	2.3375
Observations (or Sum Wgts)	8

Sorted Parameter Estimates

Term	**Estimate**	**Relative Std Error**	**Pseudo t-Ratio**	**Pseudo t-Ratio**	**Pseudo p-Value**
Long-term interest rate[Low]	0.275	0.353553	1.22		0.3307
Account-opening fee[No]	0.255	0.353553	1.13		0.3600
Initial interest[Current]	-0.17	0.353553	-0.76		0.5188
Annual fee[Current]	-0.15	0.353553	-0.67		0.5649
Annual fee[Current]*Long-term interest rate[Low]	-0.0775	0.353553	-0.34		0.7591
Annual fee[Current]*Account-opening fee[No]	-0.0725	0.353553	-0.32		0.7739
Annual fee[Current]*Initial interest[Current]	0.0525	0.353553	0.23		0.8344

No error degrees of freedom, so ordinary tests uncomputable. Relative Std Error corresponds to residual standard error of 1. Pseudo t-Ratio and p-Value calculated using Lenth PSE = 0.225 and DFE=2.3333

Prediction Profiler



8.7. Continuation of Problem 8.6. In Example 6.6, we found that all four main effects and the two-factor *AB* interaction were significant. Show that if the alternate fraction (*I=-ABCD*) is added to the 2^{4-1} design in Problem 8.6 that the analysis from the combined design produced results identical to those found in Example 6.6

The estimates for the effects from the alternate fraction are shown below. By combining the estimates from Problem 8.6 with the estimates below, the same estimates as those shown in Example 6.6 are found. For example:

$$\frac{1}{2}\left([A]+[A']\right)=\frac{1}{2}\left(-0.15+(-0.2575)\right)=-0.20375\rightarrow A$$
$$\frac{1}{2}\left([AB]+[AB']\right)=\frac{1}{2}\left(-0.0725+(-0.23)\right)=-0.15125\rightarrow AB$$
$$\frac{1}{2}\left([AB]-[AB']\right)=\frac{1}{2}\left(-0.0725-(-0.23)\right)=0.07875\rightarrow CD$$

The analysis can also be performed as full factorial in two blocks with the block effect confounded with the *ABCD* interaction. This JMP analysis for the full factorial in two blocks is also shown below. The effect estimates are the same as those shown in Example 6.6.

JMP Output
Response Response rate
Summary of Fit

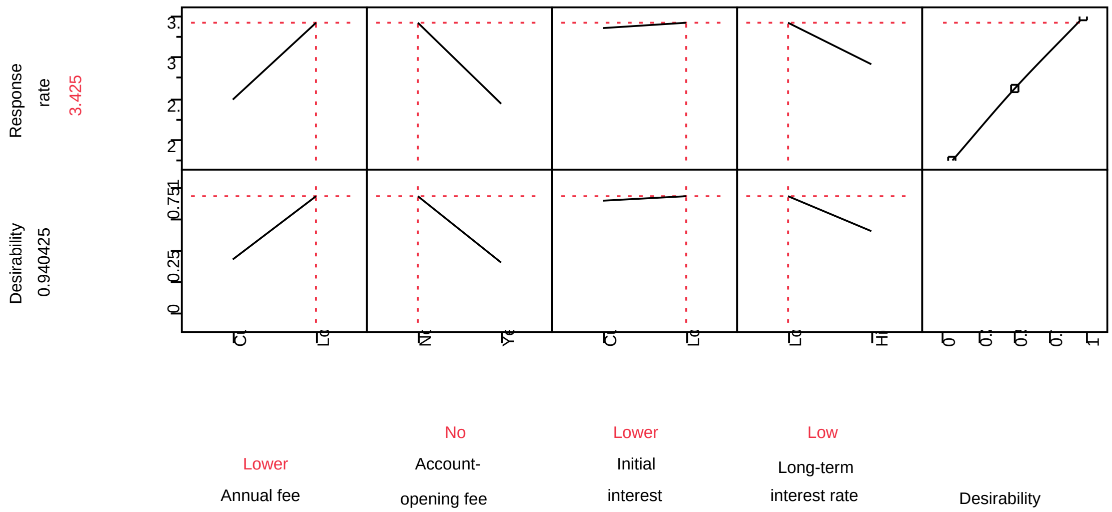
RSquare	1
RSquare Adj	.
Root Mean Square Error	.
Mean of Response	2.39
Observations (or Sum Wgts)	8

Sorted Parameter Estimates

Term	**Estimate**	**Relative Std Error**	**Pseudo t-Ratio**	**Pseudo t-Ratio**	**Pseudo p-Value**
Account-opening fee[No]	0.2625	0.353553	0.79		0.5035
Annual fee[Current]	-0.2575	0.353553	-0.77		0.5109
Annual fee[Current]*Account-opening fee[No]	-0.23	0.353553	-0.69		0.5529
Long-term interest rate[Low]	0.2225	0.353553	0.67		0.5649
Initial interest[Current]	-0.0825	0.353553	-0.25		0.8248
Annual fee[Current]*Initial interest[Current]	-0.05	0.353553	-0.15		0.8929
Annual fee[Current]*Long-term interest rate[Low]	-0.03	0.353553	-0.09		0.9355

No error degrees of freedom, so ordinary tests uncomputable. Relative Std Error corresponds to residual standard error of 1. Pseudo t-Ratio and p-Value calculated using Lenth PSE = 0.33375 and DFE=2.3333

Prediction Profiler



JMP Output
Response Response rate
Summary of Fit

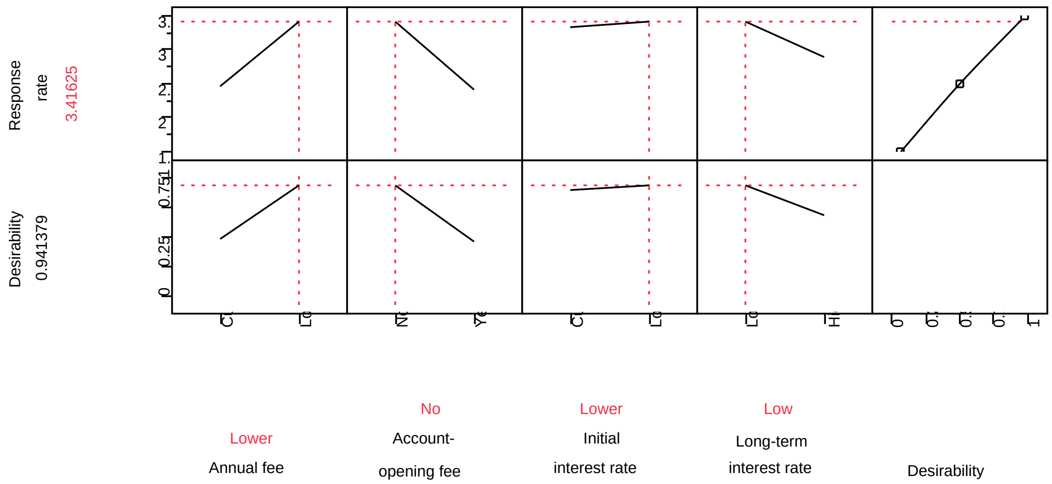
RSquare	1
RSquare Adj	.
Root Mean Square Error	.
Mean of Response	2.36375
Observations (or Sum Wgts)	16

Sorted Parameter Estimates

Term	Estimate	Relative Std Error	Pseudo t-Ratio	Pseudo t-Ratio	Pseudo p-Value
Account-opening fee[No]	0.25875	0.25	3.63		0.0150*
Long-term interest rate[Low]	0.24875	0.25	3.49		0.0174*
Annual fee[Current]	-0.20375	0.25	-2.86		0.0354*
Annual fee[Current]*Account-opening fee[No]	-0.15125	0.25	-2.12		0.0872
Initial interest rate[Current]	-0.12625	0.25	-1.77		0.1366
Initial interest rate[Current]*Long-term interest rate[Low]	0.07875	0.25	1.11		0.3194
Annual fee[Current]*Long-term interest rate[Low]	-0.05375	0.25	-0.75		0.4846
Account-opening fee[No]*Initial interest rate[Current]*Long-term interest rate[Low]	0.05375	0.25	0.75		0.4846
Account-opening fee[No]*Long-term interest rate[Low]	0.05125	0.25	0.72		0.5042
Annual fee[Current]*Account-opening fee[No]*Long-term interest rate[Low]	-0.04375	0.25	-0.61		0.5661
Block[1]	-0.02625	0.25	-0.37		0.7276
Annual fee[Current]*Account-opening fee[No]*Initial interest rate[Current]	0.02625	0.25	0.37		0.7276
Account-opening fee[No]*Initial interest rate[Current]	-0.02375	0.25	-0.33		0.7524
Annual fee[Current]*Initial interest rate[Current]*Long-term interest rate[Low]	-0.00375	0.25	-0.05		0.9601
Annual fee[Current]*Initial interest rate[Current]	0.00125	0.25	0.02		0.9867

No error degrees of freedom, so ordinary tests uncomputable. Relative Std Error corresponds to residual standard error of 1. Pseudo t-Ratio and p-Value calculated using Lenth PSE = 0.07125 and DFE=5

Prediction Profiler



8.8. Continuation of Problem 8.6. Reconsider the 2^{4-1} fractional factorial design with $I=ABCD$ from Problem 8.6. Set a partial fold-over of this fraction to isolate the *AB* interaction. Select the appropriate set of responses from the full factorial data in Example 6.6 and analyze the resulting data.

The fold-over could be based on either factor *A* or *B* to isolate the *AB* interaction. The partial fold-over shown below in block 2 was based on factor *B*, with the + level chosen (Yes).

Annual Fee	Account opening fee	Initial interest Rate	Long-term interest rate	Block	Response rate
Current	No	Current	Low	1	2.45
Lower	Yes	Current	Low	1	2.29
Lower	No	Lower	Low	1	3.39
Current	Yes	Lower	Low	1	2.32
Lower	No	Current	High	1	2.24
Current	Yes	Current	High	1	1.69
Current	No	Lower	High	1	2.29
Lower	Yes	Lower	High	1	2.03
Current	Yes	Current	Low	2	2.16
Lower	Yes	Lower	Low	2	2.44
Lower	Yes	Current	High	2	1.87
Current	Yes	Lower	High	2	2.04

The analysis is shown in the JMP Output below.

JMP Output
Response Response rate
Summary of Fit

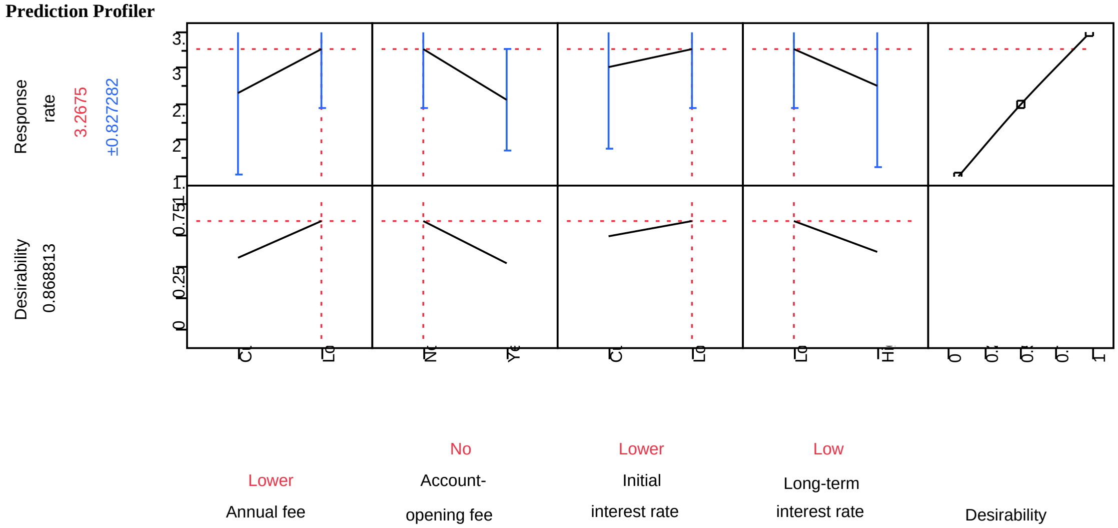
RSquare	0.956417
RSquare Adj	0.760293
Root Mean Square Error	0.205548
Mean of Response	2.2675
Observations (or Sum Wgts)	12

Sorted Parameter Estimates

Term	**Estimate**	**Std Error**	**t Ratio**	**t Ratio**	**Prob>|t|**
Long-term interest rate[Low]	0.23625	0.062936	3.75		0.0642
Account-opening fee[No]	0.255	0.072672	3.51		0.0725
Initial interest rate[Current]	-0.13625	0.062936	-2.16		0.1628
Annual fee[Current]	-0.15	0.072672	-2.06		0.1751
Annual fee[Current]*Account-opening fee[No]	-0.0975	0.072672	-1.34		0.3118
Annual fee[Current]*Long-term interest rate[Low]	-0.04375	0.062936	-0.70		0.5589

Term	Estimate	Std Error	t Ratio	t Ratio	Prob>|t|
Initial interest rate[Current]*Long-term interest rate[Low]	0.025	0.072672	0.34		0.7636
Block[1]	-0.0225	0.072672	-0.31		0.7861
Annual fee[Current]*Initial interest rate[Current]	0.01375	0.062936	0.22		0.8473

Prediction Profiler

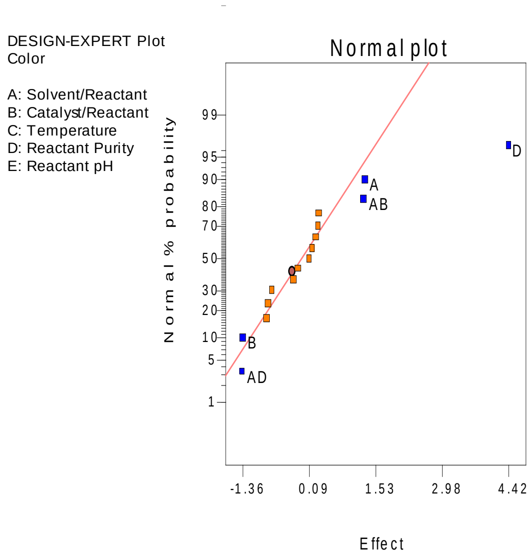


8.9. R.D. Snee ("Experimenting with a Large Number of Variables," in *Experiments in Industry: Design, Analysis and Interpretation of Results*, by R.D. Snee, L.B. Hare, and J.B. Trout, Editors, ASQC, 1985) describes an experiment in which a 2^{5-1} design with $I=ABCDE$ was used to investigate the effects of five factors on the color of a chemical product. The factors are A = solvent/reactant, B = catalyst/reactant, C = temperature, D = reactant purity, and E = reactant pH. The results obtained were as follows:

e =	-0.63	d =	6.79
a =	2.51	ade =	5.47
b =	-2.68	bde =	3.45
abe =	1.66	abd =	5.68
c =	2.06	cde =	5.22
ace =	1.22	acd =	4.38
bce =	-2.09	bcd =	4.30
abc =	1.93	$abcde$ =	4.05

(a) Prepare a normal probability plot of the effects. Which effects seem active?

Factors A, B, D, and the AB, AD interactions appear to be active.



Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	1.31	6.8644	5.98537
Model	B	-1.34	7.1824	6.26265
Error	C	-0.1475	0.087025	0.0758809
Model	D	4.42	78.1456	68.1386
Error	E	-0.8275	2.73902	2.38828
Model	AB	1.275	6.5025	5.66981
Error	AC	-0.7875	2.48062	2.16297
Model	AD	-1.355	7.3441	6.40364
Error	AE	0.3025	0.366025	0.319153
Error	BC	0.1675	0.112225	0.0978539
Error	BD	0.245	0.2401	0.209354
Error	BE	0.2875	0.330625	0.288286
Error	CD	-0.7125	2.03063	1.77059
Error	CE	-0.24	0.2304	0.200896
Error	DE	0.0875	0.030625	0.0267033
	Lenth's ME	1.95686		
	Lenth's SME	3.9727		

Design Expert Output

Response: Color

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	106.04	5	21.21	24.53	< 0.0001	significant
A	*6.86*	*1*	*6.86*	*7.94*	*0.0182*	
B	*7.18*	*1*	*7.18*	*8.31*	*0.0163*	
D	*78.15*	*1*	*78.15*	*90.37*	*< 0.0001*	
AB	*6.50*	*1*	*6.50*	*7.52*	*0.0208*	
AD	*7.34*	*1*	*7.34*	*8.49*	*0.0155*	
Residual	8.65	10	0.86			
Cor Total	114.69	15				

The Model F-value of 24.53 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.93	R-Squared	0.9246
Mean	2.71	Adj R-Squared	0.8869
C.V.	34.35	Pred R-Squared	0.8070
PRESS	22.14	Adeq Precision	14.734

	Coefficient		**Standard**	**95% CI**	**95% CI**

Factor	Estimate	DF	Error	Low	High	VIF
Intercept	2.71	1	0.23	2.19	3.23	
A-Solvent/Reactant	0.66	1	0.23	0.14	1.17	1.00
B-Catalyst/Reactant	-0.67	1	0.23	-1.19	-0.15	1.00
D-Reactant Purity	2.21	1	0.23	1.69	2.73	1.00
AB	0.64	1	0.23	0.12	1.16	1.00
AD	-0.68	1	0.23	-1.20	-0.16	1.00

Final Equation in Terms of Coded Factors:

```
Color  =
+2.71
+0.66   * A
-0.67   * B
+2.21   * D
+0.64   * A * B
-0.68   * A * D
```

Final Equation in Terms of Actual Factors:

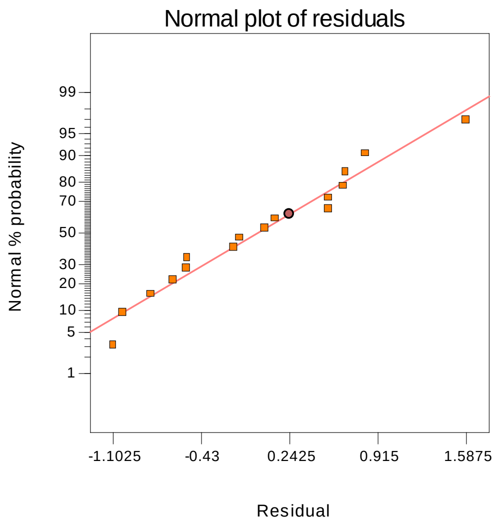
```
Color  =
+2.70750
+0.65500   * Solvent/Reactant
-0.67000   * Catalyst/Reactant
+2.21000   * Reactant Purity
+0.63750   * Solvent/Reactant * Catalyst/Reactant
-0.67750   * Solvent/Reactant * Reactant Purity
```

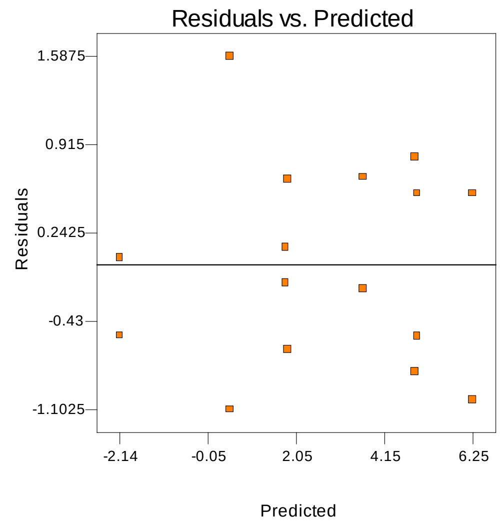
(b) Calculate the residuals. Construct a normal probability plot of the residuals and plot the residuals versus the fitted values. Comment on the plots.

Design Expert Output

Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t	Run Order
1	-0.63	0.47	-1.10	0.375	-1.500	0.225	-1.616	2
2	2.51	1.86	0.65	0.375	0.881	0.078	0.870	6
3	-2.68	-2.14	-0.54	0.375	-0.731	0.053	-0.713	14
4	1.66	1.80	-0.14	0.375	-0.187	0.003	-0.178	11
5	2.06	0.47	1.59	0.375	2.159	0.466	2.804	8
6	1.22	1.86	-0.64	0.375	-0.874	0.076	-0.863	15
7	-2.09	-2.14	0.053	0.375	0.071	0.001	0.068	10
8	1.93	1.80	0.13	0.375	0.180	0.003	0.171	3
9	6.79	6.25	0.54	0.375	0.738	0.054	0.720	4
10	5.47	4.93	0.54	0.375	0.738	0.054	0.720	5
11	3.45	3.63	-0.18	0.375	-0.248	0.006	-0.236	16
12	5.68	4.86	0.82	0.375	1.112	0.124	1.127	12
13	5.22	6.25	-1.03	0.375	-1.398	0.195	-1.478	9
14	4.38	4.93	-0.55	0.375	-0.745	0.055	-0.727	1
15	4.30	3.63	0.67	0.375	0.908	0.082	0.899	13
16	4.05	4.86	-0.81	0.375	-1.105	0.122	-1.119	7





The residual plots are satisfactory.

(c) If any factors are negligible, collapse the 2^{5-1} design into a full factorial in the active factors. Comment on the resulting design, and interpret the results.

The design becomes two replicates of a 2^3 in the factors *A*, *B* and *D*. When re-analyzing the data in three factors, *D* becomes labeled as *C*.

Design Expert Output

Response: Color
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	106.51	7	15.22	14.89	0.0005	significant
A	*6.86*	*1*	*6.86*	*6.72*	*0.0320*	
B	*7.18*	*1*	*7.18*	*7.03*	*0.0292*	
C	*78.15*	*1*	*78.15*	*76.46*	*< 0.0001*	
AB	*6.50*	*1*	*6.50*	*6.36*	*0.0357*	
AC	*7.34*	*1*	*7.34*	*7.19*	*0.0279*	
BC	*0.24*	*1*	*0.24*	*0.23*	*0.6409*	
ABC	*0.23*	*1*	*0.23*	*0.23*	*0.6476*	
Residual	8.18	8	1.02			
Lack of Fit	*0.000*	*0*				
Pure Error	*8.18*	*8*	*1.02*			
Cor Total	114.69	15				

The Model F-value of 14.89 implies the model is significant. There is only a 0.05% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.01	R-Squared	0.9287
Mean	2.71	Adj R-Squared	0.8663
C.V.	37.34	Pred R-Squared	0.7148
PRESS	32.71	Adeq Precision	11.736

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	2.71	1	0.25	2.12	3.29	
A-Solvent/Reactant	0.66	1	0.25	0.072	1.24	1.00
B-Catalyst/Reactant	-0.67	1	0.25	-1.25	-0.087	1.00
C-Reactant Purity	2.21	1	0.25	1.63	2.79	1.00
AB	0.64	1	0.25	0.055	1.22	1.00
AC	-0.68	1	0.25	-1.26	-0.095	1.00

BC	0.12	1	0.25	-0.46	0.71	1.00
ABC	-0.12	1	0.25	-0.70	0.46	1.00

Final Equation in Terms of Coded Factors:

```
Color  =
+2.71
+0.66   * A
-0.67   * B
+2.21   * C
+0.64   * A * B
-0.68   * A * C
+0.12   * B * C
-0.12   * A * B * C
```

Final Equation in Terms of Actual Factors:

```
Color  =
+2.70750
+0.65500   * Solvent/Reactant
-0.67000   * Catalyst/Reactant
+2.21000   * Reactant Purity
+0.63750   * Solvent/Reactant * Catalyst/Reactant
-0.67750   * Solvent/Reactant * Reactant Purity
+0.12250   * Catalyst/Reactant * Reactant Purity
-0.12000   * Solvent/Reactant * Catalyst/Reactant * Reactant Purity
```

8.10. An article by J.J. Pignatiello, Jr. And J.S. Ramberg in the *Journal of Quality Technology*, (Vol. 17, 1985, pp. 198-206) describes the use of a replicated fractional factorial to investigate the effects of five factors on the free height of leaf springs used in an automotive application. The factors are A = furnace temperature, B = heating time, C = transfer time, D = hold down time, and E = quench oil temperature. The data are shown in Table P81.

Table P8.1

A	*B*	*C*	*D*	*E*	Free Height		
-	-	-	-	-	7.78	7.78	7.81
+	-	-	+	-	8.15	8.18	7.88
-	+	-	+	-	7.50	7.56	7.50
+	+	-	-	-	7.59	7.56	7.75
-	-	+	+	-	7.54	8.00	7.88
+	-	+	-	-	7.69	8.09	8.06
-	+	+	-	-	7.56	7.52	7.44
+	+	+	+	-	7.56	7.81	7.69
-	-	-	-	+	7.50	7.25	7.12
+	-	-	+	+	7.88	7.88	7.44
-	+	-	+	+	7.50	7.56	7.50
+	+	-	-	+	7.63	7.75	7.56
-	-	+	+	+	7.32	7.44	7.44
+	-	+	-	+	7.56	7.69	7.62
-	+	+	-	+	7.18	7.18	7.25
+	+	+	+	+	7.81	7.50	7.59

(a) Write out the alias structure for this design. What is the resolution of this design?

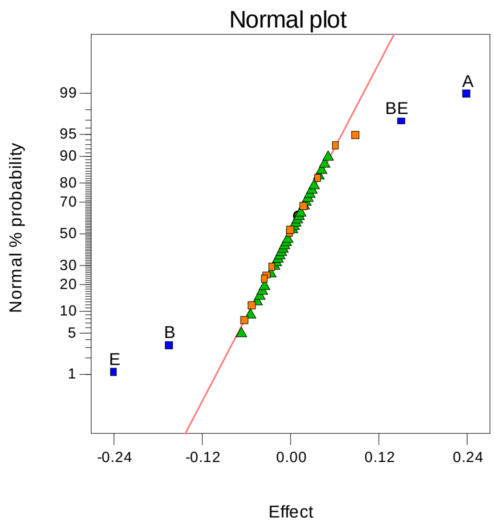
I=*ABCD*, Resolution IV

A	*(ABCD)=*	*BCD*
B	*(ABCD)=*	*ACD*
C	*(ABCD)=*	*ABD*
D	*(ABCD)=*	*ABC*
E	*(ABCD)=*	*ABCDE*
AB	*(ABCD)=*	*CD*
AC	*(ABCD)=*	*BD*
AD	*(ABCD)=*	*BC*
AE	*(ABCD)=*	*BCDE*
BE	*(ABCD)=*	*ACDE*
CE	*(ABCD)=*	*ABDE*
DE	*(ABCD)=*	*ABCE*

(b) Analyze the data. What factors influence the mean free height?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	0.242083	0.703252	24.3274
Model	B	-0.16375	0.321769	11.1309
Model	C	-0.0495833	0.0295021	1.02056
Model	D	0.09125	0.0999188	3.45646
Model	E	-0.23875	0.684019	23.6621
Model	AB	-0.0295833	0.0105021	0.363296
Model	AC	0.00125	1.875E-005	0.000648614
Model	AD	-0.0229167	0.00630208	0.218006
Model	AE	0.06375	0.0487687	1.68704
Error	BC	Aliased		
Error	BD	Aliased		
Model	BE	0.152917	0.280602	9.70679
Error	CD	Aliased		
Model	CE	-0.0329167	0.0130021	0.449777
Model	DE	0.0395833	0.0188021	0.650415
Error	Pure Error		0.627067	21.6919
	Lenth's ME	0.088057		
	Lenth's SME	0.135984		



Design Expert Output

Response:Free Height
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	1.99	4	0.50	23.74	< 0.0001	significant
A	*0.70*	*1*	*0.70*	*33.56*	*< 0.0001*	
B	*0.32*	*1*	*0.32*	*15.35*	*0.0003*	
E	*0.68*	*1*	*0.68*	*32.64*	*< 0.0001*	
BE	*0.28*	*1*	*0.28*	*13.39*	*0.0007*	
Residual	0.90	43	0.021			
Lack of Fit	*0.27*	*11*	*0.025*	*1.27*	*0.2844*	*not significant*
Pure Error	*0.63*	*32*	*0.020*			
Cor Total	2.89	47				

The Model F-value of 23.74 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.14	R-Squared	0.6883
Mean	7.63	Adj R-Squared	0.6593
C.V.	1.90	Pred R-Squared	0.6116
PRESS	1.12	Adeq Precision	13.796

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	7.63	1	0.021	7.58	7.67	
A-Furnace Temp	0.12	1	0.021	0.079	0.16	1.00
B-Heating Time	-0.082	1	0.021	-0.12	-0.040	1.00
E-Quench Temp	-0.12	1	0.021	-0.16	-0.077	1.00
BE	0.076	1	0.021	0.034	0.12	1.00

Final Equation in Terms of Coded Factors:

Free Height	=
+7.63	
+0.12	* A
-0.082	* B
-0.12	* E
+0.076	* B * E

Final Equation in Terms of Actual Factors:

Free Height	=
+7.62562	
+0.12104	* Furnace Temp
-0.081875	* Heating Time
-0.11937	* Quench Temp
+0.076458	* Heating Time * Quench Temp

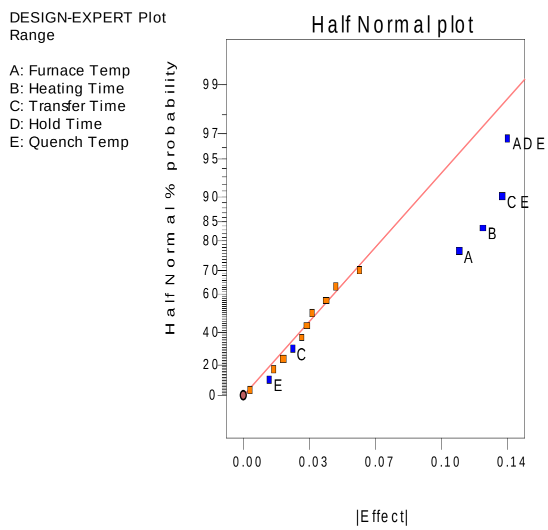
(c) Calculate the range and standard deviation of the free height for each run. Is there any indication that any of these factors affects variability in the free height?

Design Expert Output (Range)

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	0.11375	0.0517563	16.2198
Model	B	-0.12625	0.0637563	19.9804
Model	C	0.02625	0.00275625	0.863774
Error	D	0.06125	0.0150063	4.70277
Model	E	-0.01375	0.00075625	0.236999
Error	AB	0.04375	0.00765625	2.39937
Error	AC	-0.03375	0.00455625	1.42787
Error	AD	0.03625	0.00525625	1.64724
Error	AE	-0.00375	5.625E-005	0.017628
Model	BC	Aliased		
Error	BD	Aliased		
Model	BE	0.01625	0.00105625	0.331016
Error	CD	Aliased		

Model	CE	-0.13625	0.0742562	23.271
Error	DE	-0.02125	0.00180625	0.566056
Error	ABC	Aliased		
Error	ABD	Aliased		
Error	ABE	0.03125	0.00390625	1.22417
Error	ACD	Aliased		
Error	ACE	0.04875	0.00950625	2.97914
Error	ADE	0.13875	0.0770062	24.1328
Error	BCD	Aliased		
Model	BCE	Aliased		
Error	BDE	Aliased		
Error	CDE	Aliased		
	Lenth's ME	0.130136		
	Lenth's SME	0.264194		

Interaction ADE is aliased with BCE. Although the plot below identifies ADE, BCE was included in the analysis.



Design Expert Output (Range)

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	0.28	8	0.035	5.70	0.0167	significant
A	0.052	1	0.052	8.53	0.0223	
B	0.064	1	0.064	10.50	0.0142	
C	2.756E-003	1	2.756E-003	0.45	0.5220	
E	7.562E-004	1	7.562E-004	0.12	0.7345	
BC	5.256E-003	1	5.256E-003	0.87	0.3831	
BE	1.056E-003	1	1.056E-003	0.17	0.6891	
CE	0.074	1	0.074	12.23	0.0100	
BCE	0.077	1	0.077	12.69	0.0092	
Residual	0.042	7	6.071E-003			
Cor Total	0.32	15				

The Model F-value of 5.70 implies the model is significant. There is only a 1.67% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.078	R-Squared	0.8668
Mean	0.22	Adj R-Squared	0.7146
C.V.	35.52	Pred R-Squared	0.3043
PRESS	0.22	Adeq Precision	7.166

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	0.22	1	0.019	0.17	0.27	
A-Furn Temp	0.057	1	0.019	0.011	0.10	1.00
B-Heat Time	-0.063	1	0.019	-0.11	-0.017	1.00
C-Transfer Time	0.013	1	0.019	-0.033	0.059	1.00
E-Qnch Temp	-6.875E-003	1	0.019	-0.053	0.039	1.00
BC	0.018	1	0.019	-0.028	0.064	1.00
BE	8.125E-003	1	0.019	-0.038	0.054	1.00
CE	-0.068	1	0.019	-0.11	-0.022	1.00
BCE	0.069	1	0.019	0.023	0.12	1.00

Final Equation in Terms of Coded Factors:

Range	=
+0.22	
+0.057	* A
-0.063	* B
+0.013	* C
-6.875E-003	* E
+0.018	* B * C
+8.125E-003	* B * E
-0.068	* C * E
+0.069	* B * C * E

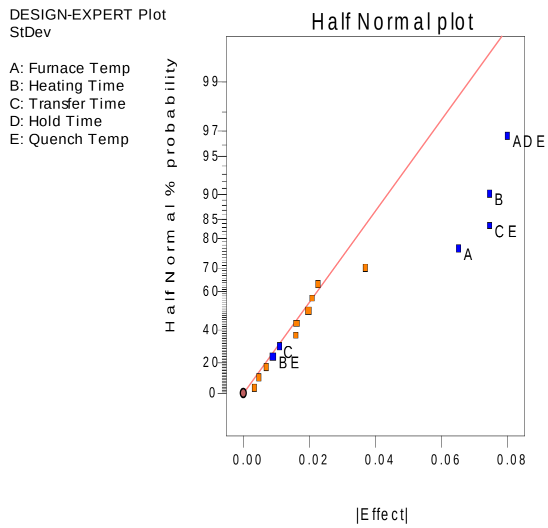
Final Equation in Terms of Actual Factors:

Range	=
+0.21937	
+0.056875	* Furnace Temp
-0.063125	* Heating Time
+0.013125	* Transfer Time
-6.87500E-003	* Quench Temp
+0.018125	* Heating Time * Transfer Time
+8.12500E-003	* Heating Time * Quench Temp
-0.068125	* Transfer Time * Quench Temp
+0.069375	* Heating Time * Transfer Time * Quench Temp

Design Expert Output (StDev)

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	0.0625896	0.0156698	16.873
Model	B	-0.0714887	0.0204425	22.0121
Model	C	0.010567	0.000446646	0.48094
Error	D	0.0353616	0.00500176	5.3858
Model	E	-0.00684034	0.000187161	0.201532
Error	AB	0.0153974	0.000948317	1.02113
Error	AC	-0.0218505	0.00190978	2.05641
Error	AD	0.0190608	0.00145326	1.56484
Error	AE	-0.00329035	4.33057E-005	0.0466308
Model	BC	Aliased		
Error	BD	Aliased		
Model	BE	0.0087666	0.000307413	0.331017
Error	CD	Aliased		
Model	CE	-0.0714816	0.0204385	22.0078
Error	DE	-0.00467792	8.75317E-005	0.0942525
Error	ABC	Aliased		
Error	ABD	Aliased		
Error	ABE	0.0155599	0.000968437	1.0428
Error	ACD	Aliased		
Error	ACE	0.0199742	0.00159587	1.7184
Error	ADE	Aliased		
Error	BCD	Aliased		
Model	BCE	0.0764346	0.023369	25.1633
Error	BDE	Aliased		
Error	CDE	Aliased		
	Lenth's ME	0.0596836		
	Lenth's SME	0.121166		

Interaction ADE is aliased with BCE. Although the plot below identifies ADE, BCE was included in the analysis.



Design Expert Output (StDev)

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	0.082	8	0.010	6.82	0.0101	significant
A	*0.016*	*1*	*0.016*	*10.39*	*0.0146*	
B	*0.020*	*1*	*0.020*	*13.56*	*0.0078*	
C	*4.466E-004*	*1*	*4.466E-004*	*0.30*	*0.6032*	
E	*1.872E-004*	*1*	*1.872E-004*	*0.12*	*0.7350*	
BC	*1.453E-003*	*1*	*1.453E-003*	*0.96*	*0.3589*	
BE	*3.074E-004*	*1*	*3.074E-004*	*0.20*	*0.6653*	
CE	*0.020*	*1*	*0.020*	*13.55*	*0.0078*	
BCE	*0.023*	*1*	*0.023*	*15.50*	*0.0056*	
Residual	0.011	7	1.508E-003			
Cor Total	0.093	15				

The Model F-value of 6.82 implies the model is significant. There is only
a 1.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.039	R-Squared	0.8863
Mean	0.12	Adj R-Squared	0.7565
C.V.	33.07	Pred R-Squared	0.4062
PRESS	0.055	Adeq Precision	7.826

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.12	1	9.708E-003	0.094	0.14	
A-Furnace Temp	0.031	1	9.708E-003	8.340E-003	0.054	1.00
B-Heating Time	-0.036	1	9.708E-003	-0.059	-0.013	1.00
C-Transfer Time	5.283E-003	1	9.708E-003	-0.018	0.028	1.00
E-Quench Temp	-3.420E-003	1	9.708E-003	-0.026	0.020	1.00
BC	9.530E-003	1	9.708E-003	-0.013	0.032	1.00
BE	4.383E-003	1	9.708E-003	-0.019	0.027	1.00
CE	-0.036	1	9.708E-003	-0.059	-0.013	1.00
BCE	0.038	1	9.708E-003	0.015	0.061	1.00

Final Equation in Terms of Coded Factors:

StDev =
+0.12

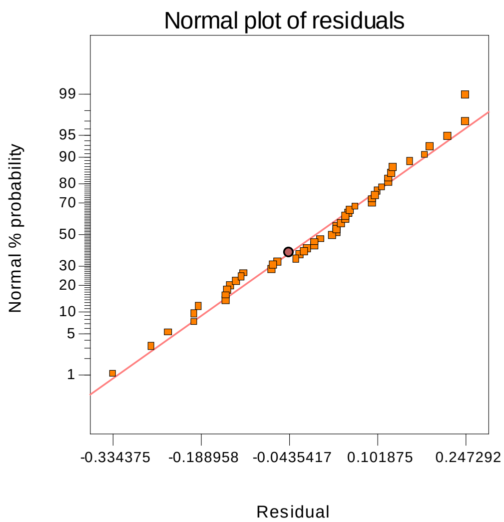
+0.031	* A
-0.036	* B
+5.283E-003	* C
-3.420E-003	* E
+9.530E-003	* B * C
+4.383E-003	* B * E
-0.036	* C * E
+0.038	* B * C * E

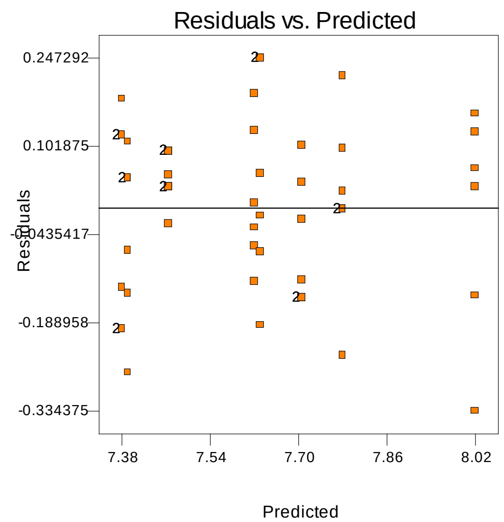
Final Equation in Terms of Actual Factors:

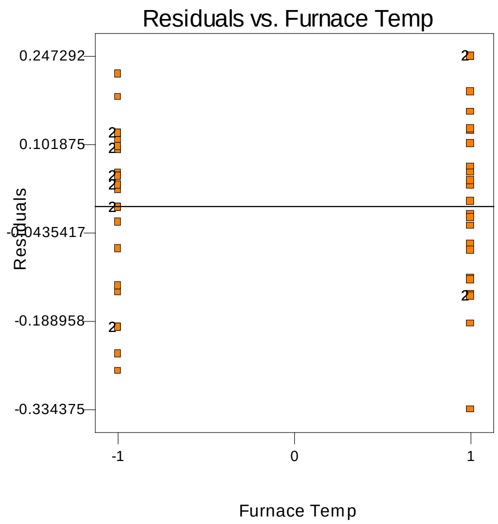
StDev	=
+0.11744	
+0.031295	* Furnace Temp
-0.035744	* Heating Time
+5.28350E-003	* Transfer Time
-3.42017E-003	* Quench Temp
+9.53040E-003	* Heating Time * Transfer Time
+4.38330E-003	* Heating Time * Quench Temp
-0.035741	* Transfer Time * Quench Temp
+0.038217	* Heating Time * Transfer Time * Quench Temp

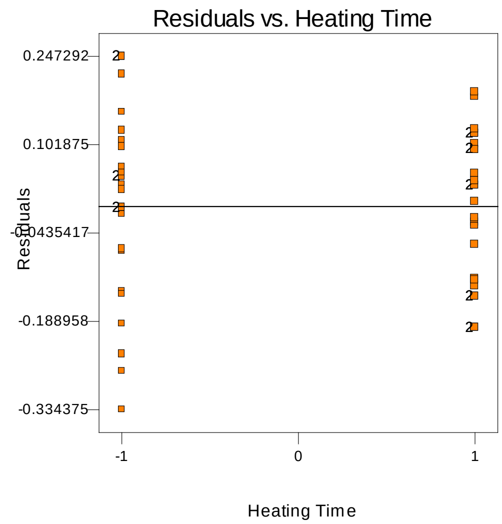
(d) Analyze the residuals from this experiment, and comment on your findings.

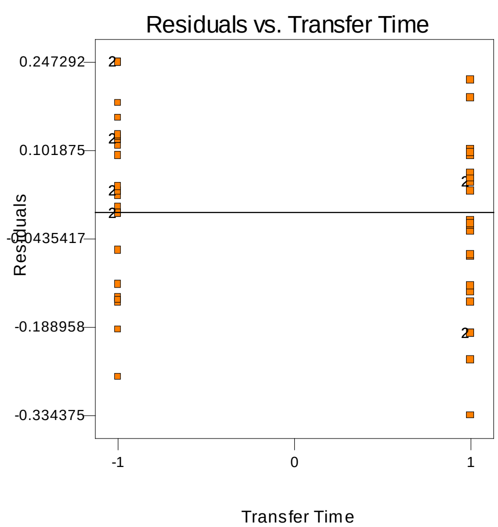
The residual plot follows. All plots are satisfactory.

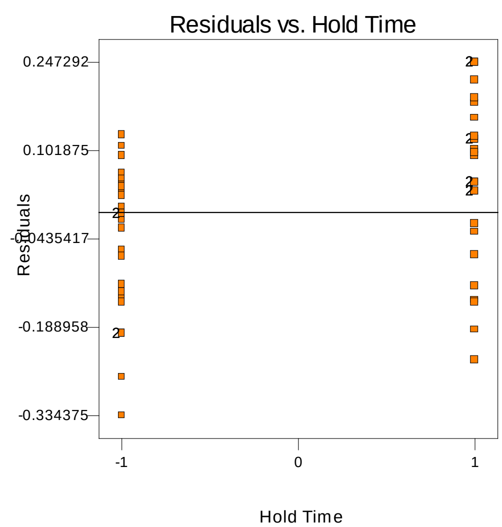


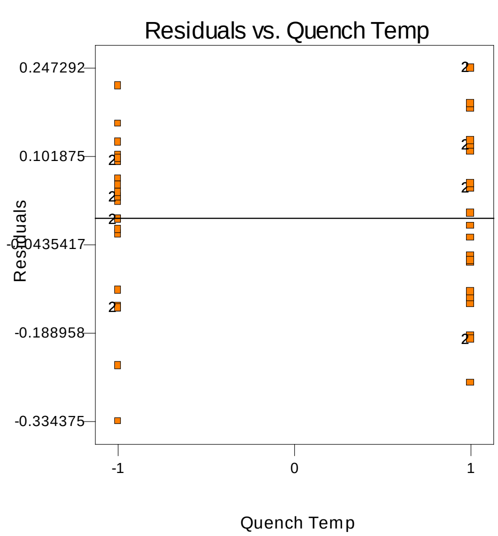



Residuals vs. Furnace Temp
0.247292
0.101875
-0.0435417
-0.188958
-0.334375
Residuals
-1
0
1
Furnace Temp


Residuals vs. Heating Time
0.247292
0.101875
-0.0435417
-0.188958
-0.334375
Residuals
-1
0
1
Heating Time


Residuals vs. Transfer Time
0.247292
0.101875
-0.0435417
-0.188958
-0.334375
Residuals
-1
0
1
Transfer Time


Residuals vs. Hold Time
0.247292
0.101875
-0.0435417
-0.188958
-0.334375
Residuals
-1
0
1
Hold Time


Residuals vs. Quench Temp
0.247292
0.101875
-0.0435417
-0.188958
-0.334375
Residuals
-1
0
1
Quench Temp

(e) Is this the best possible design for five factors in 16 runs? Specifically, can you find a fractional design for five factors in 16 runs with a higher resolution than this one?

This was not the best design. A resolution V design is possible by setting the generator equal to the highest order interaction, *ABCDE*.

8.11. An article in *Industrial and Engineering Chemistry* ("More on Planning Experiments to Increase Research Efficiency," 1970, pp. 60-65) uses a 2^{5-2} design to investigate the effect of A = condensation, B = amount of material 1, C = solvent volume, D = condensation time, and E = amount of material 2 on yield. The results obtained are as follows:

e =	23.2	*ad* =	16.9	*cd* =	23.8	*bde* =	16.8
ab =	15.5	*bc* =	16.2	*ace* =	23.4	*abcde* =	18.1

(a) Verify that the design generators used were I = *ACE* and I = *BDE*.

A	*B*	*C*	*D=BE*	*E=AC*	
-	-	-	-	+	*e*
+	-	-	+	-	*ad*
-	+	-	+	+	*bde*
+	+	-	-	-	*ab*
-	-	+	+	-	*cd*
+	-	+	-	+	*ace*
-	+	+	-	-	*bc*
+	+	+	+	+	*abcde*

(b) Write down the complete defining relation and the aliases for this design.

I=*BDE*=*ACE*=*ABCD*.

A	*(BDE)*	*=ABDE*	*A*	*(ACE)*	*=CE*	*A*	*(ABCD)*	*=BCD*	*A=ABDE=CE=BCD*
B	*(BDE)*	*=DE*	*B*	*(ACE)*	*=ABCE*	*B*	*(ABCD)*	*=ACD*	*B=DE=ABCE=ACD*
C	*(BDE)*	*=BCDE*	*C*	*(ACE)*	*=AE*	*C*	*(ABCD)*	*=ABD*	*C=BCDE=AE=ABD*
D	*(BDE)*	*=BE*	*D*	*(ACE)*	*=ACDE*	*D*	*(ABCD)*	*=ABC*	*D=BE=ACDE=ABC*
E	*(BDE)*	*=BD*	*E*	*(ACE)*	*=AC*	*E*	*(ABCD)*	*=ABCDE*	*E=BD=AC=ABCDE*
AB	*(BDE)*	*=ADE*	*AB*	*(ACE)*	*=BCE*	*AB*	*(ABCD)*	*=CD*	*AB=ADE=BCE=CD*
AD	*(BDE)*	*=ABE*	*AD*	*(ACE)*	*=CDE*	*AD*	*(ABCD)*	*=BC*	*AD=ABE=CDE=BC*

(c) Estimate the main effects.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	-1.525	4.65125	5.1831
Model	B	-5.175	53.5613	59.6858
Model	C	2.275	10.3512	11.5349
Model	D	-0.675	0.91125	1.01545
Model	E	2.275	10.3513	11.5349

(d) Prepare an analysis of variance table. Verify that the *AB* and *AD* interactions are available to use as error.

The analysis of variance table is shown below. Part (b) shows that *AB* and *AD* are aliased with other factors. If all two-factor and three factor interactions are negligible, then *AB* and *AD* could be pooled as an estimate of error.

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	79.83	5	15.97	3.22	0.2537	not significant
A	*4.65*	*1*	*4.65*	*0.94*	*0.4349*	
B	*53.56*	*1*	*53.56*	*10.81*	*0.0814*	
C	*10.35*	*1*	*10.35*	*2.09*	*0.2853*	
D	*0.91*	*1*	*0.91*	*0.18*	*0.7098*	
E	*10.35*	*1*	*10.35*	*2.09*	*0.2853*	
Residual	9.91	2	4.96			
Cor Total	89.74	7				

The "Model F-value" of 3.22 implies the model is not significant relative to the noise. There is a 25.37 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	2.23	R-Squared	0.8895
Mean	19.24	Adj R-Squared	0.6134
C.V.	11.57	Pred R-Squared	-0.7674
PRESS	158.60	Adeq Precision	5.044

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	19.24	1	0.79	15.85	22.62	
A-Condensation	-0.76	1	0.79	-4.15	2.62	1.00
B-Material 1	-2.59	1	0.79	-5.97	0.80	1.00
C-Solvent	1.14	1	0.79	-2.25	4.52	1.00
D-Time	-0.34	1	0.79	-3.72	3.05	1.00
E-Material 2	1.14	1	0.79	-2.25	4.52	1.00

Final Equation in Terms of Coded Factors:

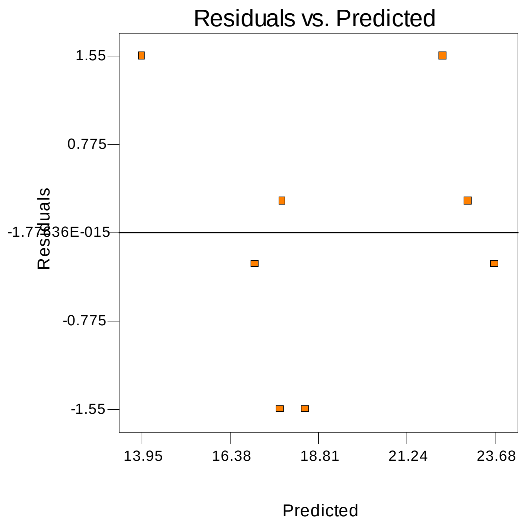
Yield	=
+19.24	
-0.76	* A
-2.59	* B
+1.14	* C
-0.34	* D
+1.14	* E

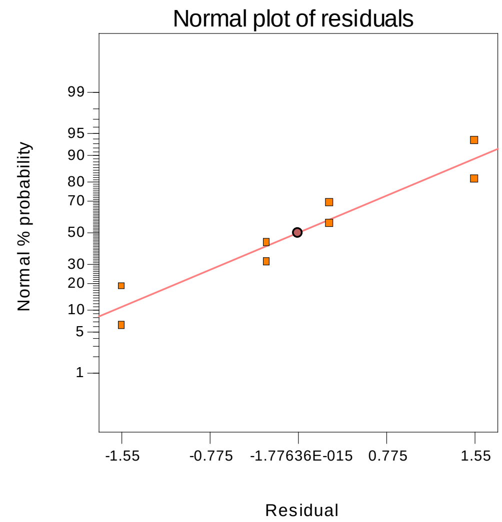
Final Equation in Terms of Actual Factors:

Yield	=
+19.23750	
-0.76250	* Condensation
-2.58750	* Material 1
+1.13750	* Solvent
-0.33750	* Time
+1.13750	* Material 2

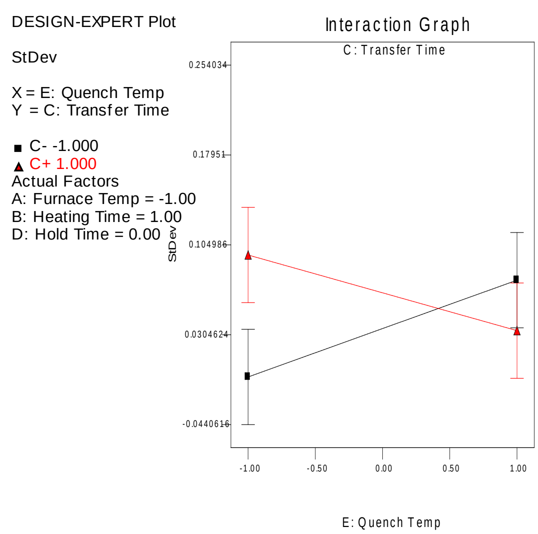
(e) Plot the residuals versus the fitted values. Also construct a normal probability plot of the residuals. Comment on the results.

The residual plots are satisfactory.





8.12. Consider the leaf spring experiment in Problem 8.10. Suppose that factor *E* (quench oil temperature) is very difficult to control during manufacturing. Where would you set factors *A*, *B*, *C* and *D* to reduce variability in the free height as much as possible regardless of the quench oil temperature used?



Run the process with *A* at the low level, *B* at the high level, *C* at the low level and *D* at either level (the low level of *D* may give a faster process).

8.13. Construct a 2^{7-2} design by choosing two four-factor interactions as the independent generators. Write down the complete alias structure for this design. Outline the analysis of variance table. What is the resolution of this design?

I=*CDEF*=*ABCG*=*ABDEFG*, Resolution IV

	A	B	C	D	E	F=CDE	G=ABC	
1	-	-	-	-	-	-	-	(1)

2	+	-	-	-	-	-	+	*ag*
3	-	+	-	-	-	-	+	*bg*
4	+	+	-	-	-	-	-	*ab*
5	-	-	+	-	-	+	+	*cfg*
6	+	-	+	-	-	+	-	*acf*
7	-	+	+	-	-	+	-	*bcf*
8	+	+	+	-	-	+	+	*abcfg*
9	-	-	-	+	-	+	-	*df*
10	+	-	-	+	-	+	+	*adfg*
11	-	+	-	+	-	+	+	*bdfg*
12	+	+	-	+	-	+	-	*abdf*
13	-	-	+	+	-	-	+	*cdg*
14	+	-	+	+	-	-	-	*acd*
15	-	+	+	+	-	-	-	*bcd*
16	+	+	+	+	-	-	+	*abcdg*
17	-	-	-	-	+	+	-	*ef*
18	+	-	-	-	+	+	+	*aefg*
19	-	+	-	-	+	+	+	*befg*
20	+	+	-	-	+	+	-	*abef*
21	-	-	+	-	+	-	+	*ceg*
22	+	-	+	-	+	-	-	*ace*
23	-	+	+	-	+	-	-	*bce*
24	+	+	+	-	+	-	+	*abceg*
25	-	-	-	+	+	-	-	*de*
26	+	-	-	+	+	-	+	*adeg*
27	-	+	-	+	+	-	+	*bdeg*
28	+	+	-	+	+	-	-	*abde*
29	-	-	+	+	+	+	+	*cdefg*
30	+	-	+	+	+	+	-	*acdef*
31	-	+	+	+	+	+	-	*bcdef*
32	+	+	+	+	+	+	+	*abcdefg*

Alias Structure

A(CDEF)= ACDEF	*A(ABCG)= BCG*	*A(ABDEFG)= BDEFG*	*A=ACDEF=BCG=BDEFG*
B(CDEF)= BCDEF	*B(ABCG)= ACG*	*B(ABDEFG)= ADEFG*	*B=BCDEF=ACG=ADEFG*
C(CDEF)= DEF	*C(ABCG)= ABG*	*C(ABDEFG)= ABCDEFG*	*C=DEF=ABG=ABCDEFG*
D(CDEF)= CEF	*D(ABCG)= ABCDG*	*D(ABDEFG)= ABEFG*	*D=CEF=ABCDG=ABEFG*
E(CDEF)= CDF	*E(ABCG)= ABCEG*	*E(ABDEFG)= ABDFG*	*E=CDF=ABCEG=ABDFG*
F(CDEF)= CDE	*F(ABCG)= ABCFG*	*F(ABDEFG)= ABDEG*	*F=CDE=ABCFG=ABDEG*
G(CDEF)= CDEFG	*G(ABCG)= ABC*	*G(ABDEFG)= ABDEF*	*G=CDEFG=ABC=ABDEF*
AB(CDEF)= ABCDEF	*AB(ABCG)= CG*	*AB(ABDEFG)= DEFG*	*AB=ABCDEF=CG=DEFG*
AC(CDEF)= ADEF	*AC(ABCG)= BG*	*AC(ABDEFG)= BCDEFG*	*AC=ADEF=BG=BCDEFG*
AD(CDEF)= ACEF	*AD(ABCG)= BCDG*	*AD(ABDEFG)= BEFG*	*AD=ACEF=BCDG=BEFG*
AE(CDEF)= ACDF	*AE(ABCG)= BCEG*	*AE(ABDEFG)= BDFG*	*AE=ACDF=BCEG=BDFG*
AF(CDEF)= ACDE	*AF(ABCG)= BCFG*	*AF(ABDEFG)= BDEG*	*AF=ACDE=BCFG=BDEG*
AG(CDEF)= ACDEFG	*AG(ABCG)= BC*	*AG(ABDEFG)= BDEF*	*AG=ACDEFG=BC=BDEF*
BD(CDEF)= BCEF	*BD(ABCG)= ACDG*	*BD(ABDEFG)= AEFG*	*BD=BCEF=ACDG=AEFG*
BE(CDEF)= BCDF	*BE(ABCG)= ACEG*	*BE(ABDEFG)= ADFG*	*BE=BCDF=ACEG=ADFG*
BF(CDEF)= BCDE	*BF(ABCG)= ACFG*	*BF(ABDEFG)= ADEG*	*BF=BCDE=ACFG=ADEG*
CD(CDEF)= EF	*CD(ABCG)= ABDG*	*CD(ABDEFG)= ABCEFG*	*CD=EF=ABDG=ABCEFG*
CE(CDEF)= DF	*CE(ABCG)= ABEG*	*CE(ABDEFG)= ABCDFG*	*CE=DF=ABEG=ABCDFG*
CF(CDEF)= DE	*CF(ABCG)= ABFG*	*CF(ABDEFG)= ABCDEG*	*CF=DE=ABFG=ABCDEG*
DG(CDEF)= CEFG	*DG(ABCG)= ABCD*	*DG(ABDEFG)= ABEF*	*DG=CEFG=ABCD=ABEF*
EG(CDEF)= CDFG	*EG(ABCG)= ABCE*	*EG(ABDEFG)= ABDF*	*EG=CDFG=ABCE=ABDF*
FG(CDEF)= CDEG	*FG(ABCG)= ABCF*	*FG(ABDEFG)= ABDE*	*FG=CDEG=ABCF=ABDE*

Analysis of Variance Table

Source	Degrees of Freedom
A	1
B	1
C	1
D	1
E	1
F	1
G	1
AB=CG	1
AC=BG	1
AD	1

AE	1
AF	1
AG=BC	1
BD	1
BE	1
BF	1
CD=EF	1
CE=DF	1
CF=DE	1
DG	1
EG	1
FG	1
Error	9
Total	31

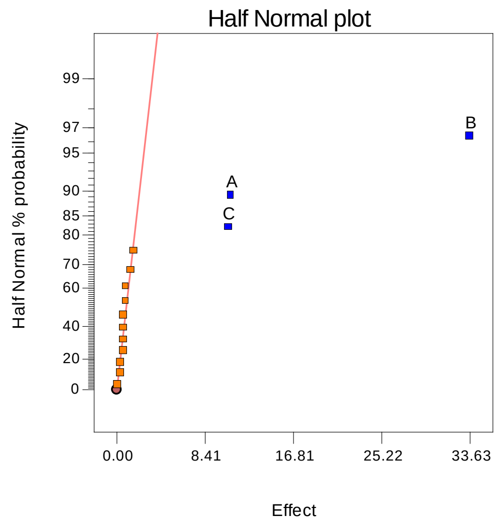
8.14. Consider the 2^5 design in Problem 6.26. Suppose that only a one-half fraction could be run. Furthermore, two days were required to take the 16 observations, and it was necessary to confound the 2^{5-1} design in two blocks. Construct the design and analyze the data.

A	*B*	*C*	*D*	*E=ABCD*		Data	Blocks = *AB*	Block
-	-	-	-	+	*e*	8	+	1
+	-	-	-	-	*a*	9	-	2
-	+	-	-	-	*b*	34	-	2
+	+	-	-	+	*abe*	52	+	1
-	-	+	-	-	*c*	16	+	1
+	-	+	-	+	*ace*	22	-	2
-	+	+	-	+	*bce*	45	-	2
+	+	+	-	-	*abc*	60	+	1
-	-	-	+	-	*d*	8	+	1
+	-	-	+	+	*ade*	10	-	2
-	+	-	+	+	*bde*	30	-	2
+	+	-	+	-	*abd*	50	+	1
-	-	+	+	+	*cde*	15	+	1
+	-	+	+	-	*acd*	21	-	2
-	+	+	+	-	*bcd*	44	-	2
+	+	+	+	+	*abcde*	63	+	1

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	10.875	473.063	8.6343
Model	B	33.625	4522.56	82.5455
Model	C	10.625	451.562	8.24188
Error	D	-0.625	1.5625	0.0285186
Error	E	0.375	0.5625	0.0102667
Error	AB	Aliased		
Error	AC	0.625	1.5625	0.0285186
Error	AD	0.875	3.0625	0.0558965
Error	AE	1.375	7.5625	0.13803
Error	BC	0.875	3.0625	0.0558965
Error	BD	-0.375	0.5625	0.0102667
Error	BE	0.125	0.0625	0.00114075
Error	CD	0.625	1.5625	0.0285186
Error	CE	0.625	1.5625	0.0285186
Error	DE	-1.625	10.5625	0.192786
	Lenth's ME	2.46263		
	Lenth's SME	5.0517		

The *AB* interaction in the above table is aliased with the three-factor interaction *BCD*, and is also confounded with blocks.



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	
Block	203.06	1	203.06			
Model	5447.19	3	1815.73	630.31	< 0.0001	significant
A	*473.06*	*1*	*473.06*	*164.22*	*< 0.0001*	
B	*4522.56*	*1*	*4522.56*	*1569.96*	*< 0.0001*	
C	*451.56*	*1*	*451.56*	*156.76*	*< 0.0001*	
Residual	31.69	11	2.88			
Cor Total	5681.94	15				

The Model F-value of 630.31 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.70	R-Squared	0.9942
Mean	30.44	Adj R-Squared	0.9926
C.V.	5.58	Pred R-Squared	0.9878
PRESS	67.04	Adeq Precision	58.100

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	30.44	1	0.42	29.50	31.37	
Block 1	3.56	1				
Block 2	-3.56					
A-Aperture	5.44	1	0.42	4.50	6.37	1.00
B-Exposure Time	16.81	1	0.42	15.88	17.75	1.00
C-Develop Time	5.31	1	0.42	4.38	6.25	1.00

Final Equation in Terms of Coded Factors:

Yield =
+30.44
+5.44 * A
+16.81 * B
+5.31 * C

Final Equation in Terms of Actual Factors:

Aperture small
Yield =
-1.56250
+0.84063 * Exposure Time

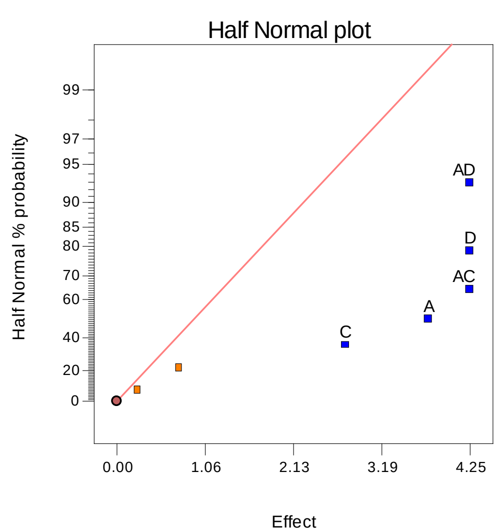
```
            +0.70833      * Develop Time

            Aperture    large
                Yield     =
            +9.31250
            +0.84063      * Exposure Time
            +0.70833      * Develop Time
```

8.15. Analyze the data in Problem 6.28 as if it came from a 2^{4-1}_{IV} design with I = *ABCD*. Project the design into a full factorial in the subset of the original four factors that appear to be significant.

Run Number	*A*	*B*	*C*	*D=ABC*		Yield (lbs)		Factor Low (-)	Levels High (+)
1	-	-	-	-	(1)	12	*A* (h)	2.5	3.0
2	+	-	-	+	*ad*	25	*B* (%)	14	18
3	-	+	-	+	*bd*	13	*C* (psi)	60	80
4	+	+	-	-	*ab*	16	*D* (°C)	225	250
5	-	-	+	+	*cd*	19			
6	+	-	+	-	*ac*	15			
7	-	+	+	-	*bc*	20			
8	+	+	+	+	*abcd*	23			

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	3.75	28.125	18.3974
Error	B	0.25	0.125	0.0817661
Model	C	2.75	15.125	9.8937
Model	D	4.25	36.125	23.6304
Error	AB	-0.75	1.125	0.735895
Model	AC	-4.25	36.125	23.6304
Model	AD	4.25	36.125	23.6304
	Lenth's ME	21.174		
	Lenth's SME	50.6734		



Design Expert Output

Response: Yield in lbs
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	151.63	5	30.32	48.52	0.0203	significant
A	*28.13*	*1*	*28.13*	*45.00*	*0.0215*	
C	*15.13*	*1*	*15.13*	*24.20*	*0.0389*	
D	*36.12*	*1*	*36.12*	*57.80*	*0.0169*	
AC	*36.12*	*1*	*36.12*	*57.80*	*0.0169*	
AD	*36.13*	*1*	*36.13*	*57.80*	*0.0169*	
Residual	1.25	2	0.62			
Cor Total	152.88	7				

The Model F-value of 48.52 implies the model is significant. There is only a 2.03% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.79	R-Squared	0.9918
Mean	17.88	Adj R-Squared	0.9714
C.V.	4.42	Pred R-Squared	0.8692
PRESS	20.00	Adeq Precision	17.892

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	17.88	1	0.28	16.67	19.08	
A-Time	1.87	1	0.28	0.67	3.08	1.00
C-Pressure	1.37	1	0.28	0.17	2.58	1.00
D-Temperature	2.13	1	0.28	0.92	3.33	1.00
AC	-2.13	1	0.28	-3.33	-0.92	1.00
AD	2.13	1	0.28	0.92	3.33	1.00

Final Equation in Terms of Coded Factors:

```
       Yield   =
      +17.88
       +1.87   * A
       +1.37   * C
       +2.13   * D
       -2.13   * A * C
       +2.13   * A * D
```

Final Equation in Terms of Actual Factors:

```
        Yield   =
   +227.75000
    -94.50000   * Time
     +2.47500   * Pressure
     -1.70000   * Temperature
     -0.85000   * Time * Pressure
     +0.68000   * Time * Temperature
```

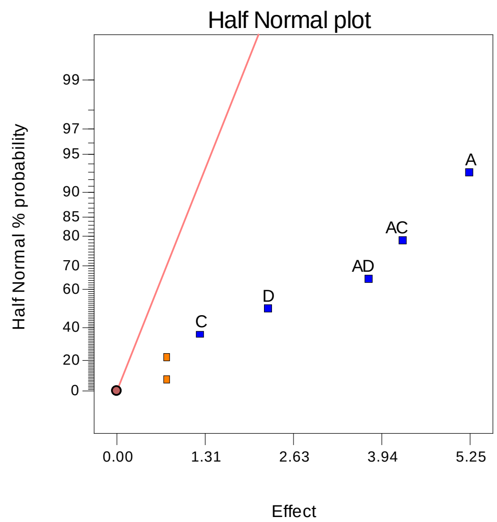
8.16. Repeat Problem 8.15 using I = -*ABCD*. Does use of the alternate fraction change your interpretation of the data?

Run Number	*A*	*B*	*C*	*D=ABC*		Yield (lbs)		Factor Low (-)	Levels High (+)
1	-	-	-	+	*d*	10	*A* (h)	2.5	3.0
2	+	-	-	-	*a*	18	*B* (%)	14	18
3	-	+	-	-	*b*	13	*C* (psi)	60	80
4	+	+	-	+	*abd*	24	*D* (°C)	225	250
5	-	-	+	-	*c*	17			
6	+	-	+	+	*acd*	21			
7	-	+	+	+	*bcd*	17			
8	+	+	+	-	*abc*	15			

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	5.25	55.125	40.8712

Error	B	0.75	1.125	0.834106
Model	C	1.25	3.125	2.31696
Model	D	2.25	10.125	7.50695
Error	AB	-0.75	1.125	0.834106
Model	AC	-4.25	36.125	26.7841
Model	AD	3.75	28.125	20.8526
	Lenth's ME	12.7044		
	Lenth's SME	30.404		



Design Expert Output

Response: Yield in lbs

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	132.63	5	26.52	23.58	0.0412	significant
A	*55.13*	*1*	*55.13*	*49.00*	*0.0198*	
C	*3.13*	*1*	*3.13*	*2.78*	*0.2375*	
D	*10.13*	*1*	*10.13*	*9.00*	*0.0955*	
AC	*36.13*	*1*	*36.13*	*32.11*	*0.0298*	
AD	*28.13*	*1*	*28.13*	*25.00*	*0.0377*	
Residual	2.25	2	1.12			
Cor Total	134.88	7				

The Model F-value of 23.58 implies the model is significant. There is only a 4.12% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.06	R-Squared	0.9833
Mean	16.88	Adj R-Squared	0.9416
C.V.	6.29	Pred R-Squared	0.7331
PRESS	36.00	Adeq Precision	14.425

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	16.88	1	0.37	15.26	18.49	
A-Time	2.63	1	0.37	1.01	4.24	1.00
C-Pressure	0.63	1	0.37	-0.99	2.24	1.00
D-Temperature	1.13	1	0.37	-0.49	2.74	1.00
AC	-2.13	1	0.37	-3.74	-0.51	1.00
AD	1.88	1	0.37	0.26	3.49	1.00

Final Equation in Terms of Coded Factors:

Yield =
+16.88

```
                +2.63     * A
                +0.63     * C
                +1.13     * D
                -2.13     * A * C
                +1.88     * A * D

     Final Equation in Terms of Actual Factors:

                Yield     =
           +190.50000
            -72.50000     * Time
             +2.40000     * Pressure
             -1.56000     * Temperature
             -0.85000     * Time * Pressure
             +0.60000     * Time * Temperature
```

8.17. Project the 2_{IV}^{4-1} design in Example 8.1 into two replicates of a 2^2 design in the factors *A* and *B*. Analyze the data and draw conclusions.

The *Design Expert* output below does not identify a significant effect.

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	728.50	3	242.83	0.41	0.7523	not significant
A	*722.00*	*1*	*722.00*	*1.23*	*0.3291*	
B	*4.50*	*1*	*4.50*	*7.682E-003*	*0.9344*	
AB	*2.00*	*1*	*2.00*	*3.414E-003*	*0.9562*	
Residual	2343.00	4	585.75			
Lack of Fit	*0.000*	*0*				
Pure Error	*2343.00*	*4*	*585.75*			
Cor Total	3071.50	7				

The "Model F-value" of 0.41 implies the model is not significant relative to the noise. There is a 75.23 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	24.20	R-Squared	0.2372
Mean	70.75	Adj R-Squared	-0.3349
C.V.	34.21	Pred R-Squared	-2.0513
PRESS	9372.00	Adeq Precision	1.198

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	70.75	1	8.56	46.99	94.51	
A-Temperature	9.50	1	8.56	-14.26	33.26	1.00
B-Pressure	0.75	1	8.56	-23.01	24.51	1.00
AB	-0.50	1	8.56	-24.26	23.26	1.00

```
     Final Equation in Terms of Coded Factors:

       Filtration Rate    =
               +70.75
                +9.50     * A
                +0.75     * B
                -0.50     * A * B

     Final Equation in Terms of Actual Factors:

       Filtration Rate    =
            +70.75000
             +9.50000     * Temperature
             +0.75000     * Pressure
             -0.50000     * Temperature * Pressure
```

8.18. Construct a 2_{III}^{5-2} design. Determine the effects that may be estimated if a full fold over of this design is performed.

The 2_{III}^{5-2} design is shown below.

A	*B*	*C*	*D=AB*	*E=AC*	
-	-	-	+	+	*de*
+	-	-	-	-	*a*
-	+	-	-	+	*be*
+	+	-	+	-	*abd*
-	-	+	+	-	*cd*
+	-	+	-	+	*ace*
-	+	+	-	-	*bc*
+	+	+	+	+	*abcde*

The design with the fold over included is as follows.

Block	*A*	*B*	*C*	*D=AB*	*E=AC*	
1	-	-	-	+	+	*de*
1	+	-	-	-	-	*a*
1	-	+	-	-	+	*be*
1	+	+	-	+	-	*abd*
1	-	-	+	+	-	*cd*
1	+	-	+	-	+	*ace*
1	-	+	+	-	-	*bc*
1	+	+	+	+	+	*abcde*
2	+	+	+	-	-	*abc*
2	-	+	+	+	+	*bcde*
2	+	-	+	+	-	*acd*
2	-	-	+	-	+	*ce*
2	+	+	-	-	+	*abe*
2	-	+	-	+	-	*bd*
2	+	-	-	+	+	*ade*
2	-	-	-	-	-	(1)

The effects are shown in the table below.

Principal Fraction	Second Fraction
$\ell_A = A+BD+CE$	$\ell^*_A = A-BD-CE$
$\ell_B = B + AD$	$\ell^*_B = B - AD$
$\ell_C = C + AE$	$\ell^*_C = C - AE$
$\ell_D = D + AB$	$\ell^*_D = D - AB$
$\ell_E = E + AC$	$\ell^*_E = E - AC$
$\ell_{BC} = BC + DE$	$\ell^*_{BC} = BC + DE$
$\ell_{BE} = BE + CD$	$\ell^*_{BE} = BE + CD$

By combining the two fractions we can estimate the following:

$(\ell_i + \ell^*_i)/2$	$(\ell_i - \ell^*_i)/2$
A	*BD+CE*
B	*AD*

C	*AE*
D	*AB*
E	*AC*
BC+DE	
BE+CD	

These estimates are confirmed with the Design Expert output shown below.

Design Expert Output

Design Matrix Evaluation for Factorial Reduced 3FI Model

Factorial Effects Aliases
[Est. Terms] Aliased Terms
[Intercept] = Intercept
[Block 1] = Block 1 + ABD + ACE
[Block 2] = Block 2 - ABD - ACE
[A] = A
[B] = B + CDE
[C] = C + BDE
[D] = D + BCE
[E] = E + BCD
[AB] = AB
[AC] = AC
[AD] = AD
[AE] = AE
[BC] = BC + DE
[BD] = BD + CE
[BE] = BE + CD
[ABC] = ABC + ADE
[ABE] = ABE + ACD

8.19. Construct a 2_{III}^{6-3} design. Determine the effects that may be estimated if a second fraction of this design is run with all signs reversed.

The 2_{III}^{6-3} design is shown below.

A	*B*	*C*	*D=AB*	*E=AC*	*F=BC*	
-	-	-	+	+	+	*def*
+	-	-	-	-	+	*af*
-	+	-	-	+	-	*be*
+	+	-	+	-	-	*abd*
-	-	+	+	-	-	*cd*
+	-	+	-	+	-	*ace*
-	+	+	-	-	+	*bcf*
+	+	+	+	+	+	*abcdef*

The effects are shown in the table below.

Principal Fraction	Second Fraction
$\ell_A = A+BD+CE$	$\ell^*_A = A-BD-CE$
$\ell_B = B+AD+CF$	$\ell^*_B = B-AD-CF$
$\ell_C = C+AE+BF$	$\ell^*_C = C-AE-BF$
$\ell_D = D+AB+EF$	$\ell^*_D = D-AB-EF$
$\ell_E = E+AC+DF$	$\ell^*_E = E-AC-DF$
$\ell_F = F+BC+DE$	$\ell^*_F = F-BC-DE$
$\ell_{BE} = BE+CD+AF$	$\ell^*_{BE} = BE+CD+AF$

By combining the two fractions we can estimate the following:

$(\ell_i + \ell^*_i)/2$	$(\ell_i - \ell^*_i)/2$
A	*BD+CE*
B	*AD+CF*
C	*AE+BF*
D	*AB+EF*
E	*AC+DF*
F	*BC+DE*
BE+CD+AF	

8.20. Consider the 2^{6-3}_{III} design in Problem 8.19. Determine the effects that may be estimated if a single factor fold over of this design is run with the signs for factor A reversed.

Principal Fraction	Second Fraction
ℓ_A=*A+BD+CE*	ℓ^*_A=*-A+BD+CE*
ℓ_B=*B+AD+CF*	ℓ^*_B=*B-AD+CF*
ℓ_C=*C+AE+BF*	ℓ^*_C=*C-AE+BF*
ℓ_D=*D+AB+EF*	ℓ^*_D=*D-AB+EF*
ℓ_E=*E+AC+DF*	ℓ^*_E=*E-AC+DF*
ℓ_F=*F+BC+DE*	ℓ^*_F=*F+BC+DE*
ℓ_{BE}=*BE+CD+AF*	ℓ^*_{BE}=*BE+CD-AF*

By combining the two fractions we can estimate the following:

$(\ell_i - \ell^*_I)/2$	$(\ell_i + \ell^*_I)/2$
A	*BD+CE*
AD	*B+CF*
AE	*C+BF*
AB	*D+EF*
AC	*E+DF*
	F+BC+DE
AF	

8.21. Fold over the 2^{7-4}_{III} design in Table 8.19 to produce a eight-factor design. Verify that the resulting design is a 2^{8-4}_{IV} design. Is this a minimal design?

	H	*A*	*B*	*C*	*D=AB*	*E=AC*	*F=BC*	*G=ABC*
	+	-	-	-	+	+	+	-
Original	+	+	-	-	-	-	+	+
Design	+	-	+	-	-	+	-	+
	+	+	+	-	+	-	-	-
	+	-	-	+	+	-	-	+
	+	+	-	+	-	+	-	-
	+	-	+	+	-	-	+	-
	+	+	+	+	+	+	+	+
	-	+	+	+	-	-	-	+
Second	-	-	+	+	+	+	-	-
Set of	-	+	-	+	+	-	+	-
Runs w/	-	-	-	+	-	+	+	+
all Signs	-	+	+	-	-	+	+	-
Switched	-	-	+	-	+	-	+	+
	-	+	-	-	+	+	-	+
	-	-	-	-	-	-	-	-

After folding the original design over, we add a new factor *H*, and we have a design with generators *D=ABH*, *E=ACH*, *F=BCH*, and *G=ABC*. This is a 2_{IV}^{8-4} design. It is a minimal design, since it contains 2*k*=2(8)=16 runs.

8.19. Fold over a 2_{III}^{5-2} design to produce a six-factor design. Verify that the resulting design is a 2_{IV}^{6-2} design. Compare this 2_{IV}^{6-2} design to the in Table 8.10.

	F	*A*	*B*	*C*	*D=AB*	*E=BC*
	+	-	-	-	+	+
Original	+	+	-	-	-	+
Design	+	-	+	-	-	-
	+	+	+	-	+	-
	+	-	-	+	+	-
	+	+	-	+	-	-
	+	-	+	+	-	+
	+	+	+	+	+	+
	-	+	+	+	-	-
Second	-	-	+	+	+	-
Set of	-	+	-	+	+	+
Runs w/	-	-	-	+	-	+
all Signs	-	+	+	-	-	+
Switched	-	-	+	-	+	+
	-	+	-	-	+	-
	-	-	-	-	-	-

If we relabel the factors from left to right as *A*, *B*, *C*, *D*, *E*, *F*, then this design becomes 2_{IV}^{6-2} with generators I=ABDF and I=BCEF. It is not a minimal design, since 2*k*=2(6)=12 runs, and the design contains 16 runs.

8.23. An industrial engineer is conducting an experiment using a Monte Carlo simulation model of an inventory system. The independent variables in her model are the order quantity (*A*), the reorder point (B), the setup cost (*C*), the backorder cost (*D*), and the carrying cost rate (*E*). The response variable is average annual cost. To conserve computer time, she decides to investigate these factors using a 2_{III}^{5-2} design with I = *ABD* and I = *BCE*. The results she obtains are *de* = 95, *ae* = 134, *b* = 158, *abd* = 190, *cd* = 92, *ac* = 187, *bce* = 155, and *abcde* = 185.

(a) Verify that the treatment combinations given are correct. Estimate the effects, assuming three-factor and higher interactions are negligible.

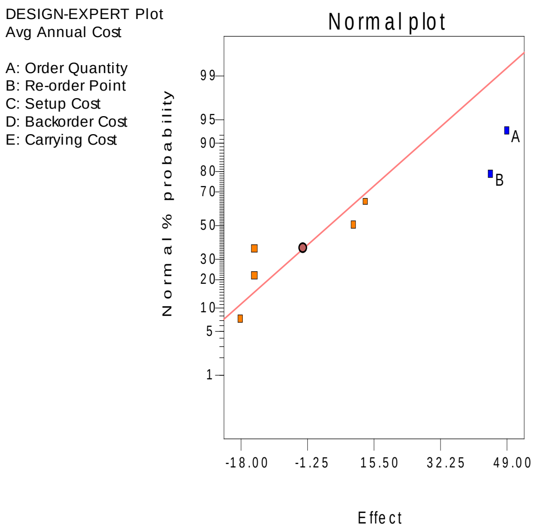
The treatment combinations given are correct. The effects are shown in the *Design Expert* output and the normal and half normal probability plot of effects identify factors *A* and *B* as important.

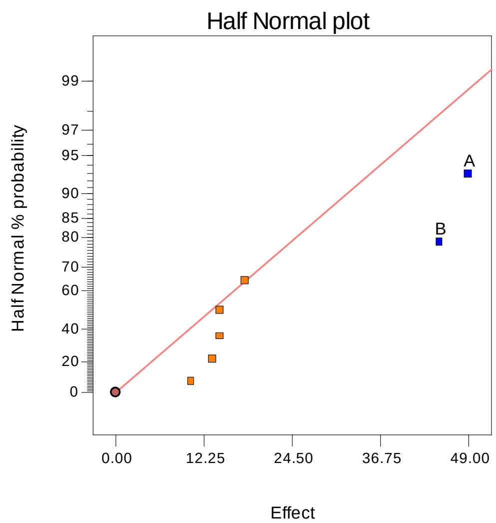
A	*B*	*C*	*D=AB*	*E=BC*	
-	-	-	+	+	*de*
+	-	-	-	+	*ae*
-	+	-	-	-	*b*
+	+	-	+	-	*abd*
-	-	+	+	-	*cd*
+	-	+	-	-	*ac*
-	+	+	-	+	*bce*
+	+	+	+	+	*abcde*

Design Expert Output

Term	Effect	SumSqr	% Contribtn

Model	Intercept			
Model	A	49	4802	43.9502
Model	B	45	4050	37.0675
Error	C	10.5	220.5	2.01812
Error	D	-18	648	5.93081
Error	E	-14.5	420.5	3.84862
Error	AC	13.5	364.5	3.33608
Error	AE	-14.5	420.5	3.84862
	Lenth's ME	81.8727		
	Lenth's SME	195.937		





(b) Suppose that a second fraction is added to the first, for example *ade* = 136, *e* = 93, *ab* = 187, *bd* = 153, *acd* = 139, *c* = 99, *abce* = 191, and *bcde* = 150. How was this second fraction obtained? Add this data to the original fraction, and estimate the effects.

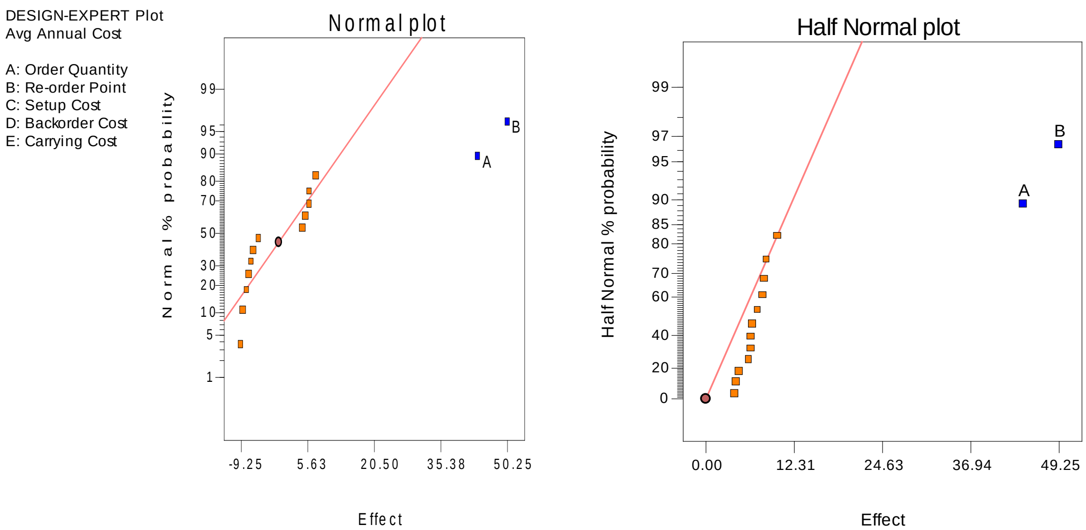
This second fraction is formed by reversing the signs of factor A.

A	*B*	*C*	*D=AB*	*E=BC*	
+	-	-	+	+	*ade*
-	-	-	-	+	*e*
+	+	-	-	-	*ab*
-	+	-	+	-	*bd*
+	-	+	+	-	*acd*
-	-	+	-	-	*c*
+	+	+	-	+	*abce*
-	+	+	+	+	*bcde*

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	44.25	7832.25	39.5289
Model	B	49.25	9702.25	48.9666
Error	C	6.5	169	0.852932
Error	D	-8	256	1.29202
Error	E	-8.25	272.25	1.37403
Error	AB	-10	400	2.01877
Error	AC	7.25	210.25	1.06112
Error	AD	-4.25	72.25	0.364641
Error	AE	-6	144	0.726759
Error	BD	4.75	90.25	0.455486

Error	CD	-8.5	289	1.45856
Error	DE	6.25	156.25	0.788584
Error	ACD	-6.25	156.25	0.788584
Error	ADE	4	64	0.323004
	Lenth's ME	25.1188		
	Lenth's SME	51.5273		



(c) Suppose that the fraction *abc* = 189, *ce* = 96, *bcd* = 154, *acde* = 135, *abe* = 193, *bde* = 152, *ad* = 137, and (1) = 98 was run. How was this fraction obtained? Add this data to the original fraction and estimate the effects.

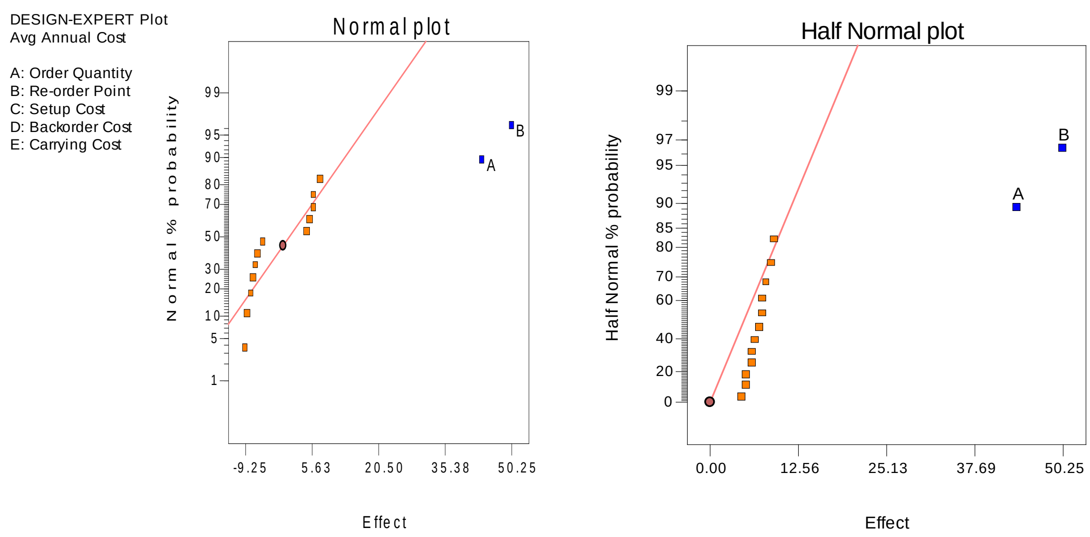
This second fraction is formed by reversing the signs of all factors.

A	*B*	*C*	*D=AB*	*E=BC*	
+	+	+	-	-	*abc*
-	+	+	+	-	*bcd*
+	-	+	+	+	*acde*
-	-	+	-	+	*ce*
+	+	-	-	+	*abe*
-	+	-	+	+	*bde*
+	-	-	+	-	*ad*
-	-	-	-	-	(1)

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	43.75	7656.25	38.1563
Model	B	50.25	10100.3	50.3364
Error	C	4.5	81	0.403678
Error	D	-8.75	306.25	1.52625
Error	E	-7.5	225	1.12133
Error	AB	-9.25	342.25	1.70566
Error	AC	6	144	0.71765
Error	AD	-5.25	110.25	0.549451
Error	AE	-6.5	169	0.842242
Error	BC	-7	196	0.976801
Error	BD	5.25	110.25	0.549451
Error	BE	6	144	0.71765
Error	ABC	-8	256	1.27582
Error	ABE	7.5	225	1.12133

Lenth's ME	26.5964
Lenth's SME	54.5583



8.24. Construct a 2^{5-1} design. Show how the design may be run in two blocks of eight observations each. Are any main effects or two-factor interactions confounded with blocks?

A	*B*	*C*	*D*	*E=ABCD*		Blocks = *AB*	Block
-	-	-	-	+	*e*	+	1
+	-	-	-	-	*a*	-	2
-	+	-	-	-	*b*	-	2
+	+	-	-	+	*abe*	+	1
-	-	+	-	-	*c*	+	1
+	-	+	-	+	*ace*	-	2
-	+	+	-	+	*bce*	-	2
+	+	+	-	-	*abc*	+	1
-	-	-	+	-	*d*	+	1
+	-	-	+	+	*ade*	-	2
-	+	-	+	+	*bde*	-	2
+	+	-	+	-	*abd*	+	1
-	-	+	+	+	*cde*	+	1
+	-	+	+	-	*acd*	-	2
-	+	+	+	-	*bcd*	-	2
+	+	+	+	+	*abcde*	+	1

Blocks are confounded with *AB* and *CDE*.

8.25. Construct a 2^{7-2} design. Show how the design may be run in four blocks of eight observations each. Are any main effects or two-factor interactions confounded with blocks?

	A	*B*	*C*	*D*	*E*	*F=CDE*	*G=ABC*		Block=ACE	Block=BFG	Block assignment
1	-	-	-	-	-	-	-	(1)	-	-	1
2	+	-	-	-	-	-	+	*ag*	+	+	4
3	-	+	-	-	-	-	+	*bg*	-	-	1
4	+	+	-	-	-	-	-	*ab*	+	+	4
5	-	-	+	-	-	+	+	*cfg*	+	-	3
6	+	-	+	-	-	+	-	*acf*	-	+	2
7	-	+	+	-	-	+	-	*bcf*	+	-	3
8	+	+	+	-	-	+	+	*abcfg*	-	+	2
9	-	-	-	+	-	+	-	*df*	-	+	2
10	+	-	-	+	-	+	+	*adfg*	+	-	3
11	-	+	-	+	-	+	+	*bdfg*	-	+	2
12	+	+	-	+	-	+	-	*abdf*	+	-	3
13	-	-	+	+	-	-	+	*cdg*	+	+	4
14	+	-	+	+	-	-	-	*acd*	-	-	1
15	-	+	+	+	-	-	-	*bcd*	+	+	4
16	+	+	+	+	-	-	+	*abcdg*	-	-	1
17	-	-	-	-	+	+	-	*ef*	+	+	4
18	+	-	-	-	+	+	+	*aefg*	-	-	1
19	-	+	-	-	+	+	+	*befg*	+	+	4
20	+	+	-	-	+	+	-	*abef*	-	-	1
21	-	-	+	-	+	-	+	*ceg*	-	+	2
22	+	-	+	-	+	-	-	*ace*	+	-	3
23	-	+	+	-	+	-	-	*bce*	-	+	2
24	+	+	+	-	+	-	+	*abceg*	+	-	3
25	-	-	-	+	+	-	-	*de*	+	-	3
26	+	-	-	+	+	-	+	*adeg*	-	+	2
27	-	+	-	+	+	-	+	*bdeg*	+	-	3
28	+	+	-	+	+	-	-	*abde*	-	+	2
29	-	-	+	+	+	+	+	*cdefg*	-	-	1
30	+	-	+	+	+	+	-	*acdef*	+	+	4
31	-	+	+	+	+	+	-	*bcdef*	-	-	1
32	+	+	+	+	+	+	+	*abcdefg*	+	+	4

Blocks are confounded with *ACE*, *BFG*, and *ABCEFG*.

8.26. ***Nonregular fractions of the 2^k*** **[John (1971)].** Consider a 2^4 design. We must estimate the four main effects and the six two-factor interactions, but the full 2^4 factorial cannot be run. The largest possible block contains 12 runs. These 12 runs can be obtained from the four one-quarter fractions defined by $I = \pm AB = \pm ACD = \pm BCD$ by omitting the principal fraction. Show how the remaining three 2^{4-2} fractions can be combined to estimate the required effects, assuming that three-factor and higher interactions are negligible. This design could be thought of as a three-quarter fraction.

The four 2^{4-2} fractions are as follows:

(1) I=+*AB*=+*ACD*=+*BCD*
Runs: *c*,*d*,*ab*,*abcd*

(2) I=+*AB*=-*ACD*=-*BCD*
Runs: (1), *cd, abc, abd*

(3) I=-*AB*=+*ACD*=-*BCD*
Runs: *a, bc, bd, acd*

(4) I=-*AB*=-*ACD*=+*BCD*
Runs: *b, ac, ad, bcd*

If we do not run the principal fraction (1), then we can combine the remaining 3 fractions to 3 one-half fractions of the 2^4 as follows:

Fraction 1: (2) + (3) implies I=-*BCD*. This fraction estimates: *A*, *AB*, *AC*, and *AD*
Fraction 2: (2) + (4) implies I=-*ACD*. This fraction estimates: *B*, *BC, BD*, and *AB*
Fraction 3: (3) + (4) implies I=-*AB*. This fraction estimates: *C, D*, and *CD*

In estimating these effects we assume that all three-factor and higher interactions are negligible. Note that *AB* is estimated in two of the one-half fractions: 1 and 2. We would average these quantities and obtain a single estimate of *AB*. John (1971, pp. 161-163) discusses this design and shows that the estimates obtained above are also the least squares estimates. John also derives the variances and covariances of these estimators.

8.27. Carbon anodes used in a smelting process are baked in a ring furnace. An experiment is run in the furnace to determine which factors influence the weight of packing material that is stuck to the anodes after baking. Six variables are of interest, each at two levels: *A* = pitch/fines ratio (0.45, 0.55); *B* = packing material type (1, 2); *C* = packing material temperature (ambient, 325 C); *D* = flue location (inside, outside); *E* = pit temperature (ambient, 195 C); and *F* = delay time before packing (zero, 24 hours). A 2^{6-3} design is run, and three replicates are obtained at each of the design points. The weight of packing material stuck to the anodes is measured in grams. The data in run order are as follows: *abd* = (984, 826, 936); *abcdef* = (1275, 976, 1457); *be* = (1217, 1201, 890); *af* = (1474, 1164, 1541); *def* = (1320, 1156, 913); *cd* = (765, 705, 821); *ace* = (1338, 1254, 1294); and *bcf* = (1325, 1299, 1253). We wish to minimize the amount stuck packing material.

(a) Verify that the eight runs correspond to a 2_{III}^{6-3} design. What is the alias structure?

A	*B*	*C*	*D=AB*	*E=AC*	*F=BC*	
-	-	-	+	+	+	*def*
+	-	-	-	-	+	*af*
-	+	-	-	+	-	*be*
+	+	-	+	-	-	*abd*
-	-	+	+	-	-	*cd*
+	-	+	-	+	-	*ace*
-	+	+	-	-	+	*bcf*
+	+	+	+	+	+	*abcdef*

I=*ABD*=*ACE*=*BCF*=*BCDE*=*ACDF*=*ABEF*=*DEF*, Resolution III

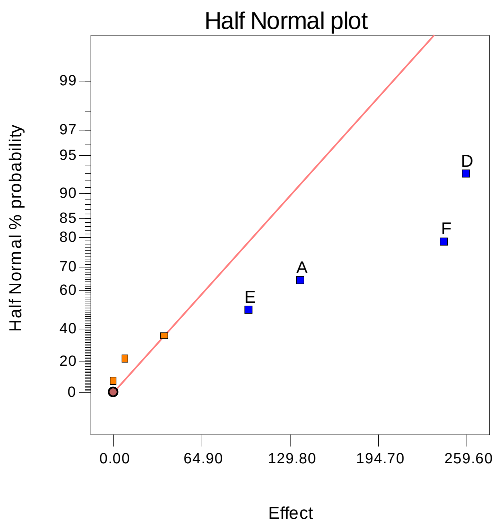
A=*BD*=*CE*=*CDF*=*BEF*
B=*AD*=*CF*=*CDE*=*AEF*
C=*AE*=*BF*=*BDE*=*ADF*
D=*AB*=*EF*=*BCE*=*ACF*
E=*AC*=*DF*=*BCD*=*ABF*
F=*BC*=*DE*=*ACD*=*ABE*
CD=*BE*=*AF*=*ABC*=*ADE*=*BDF*=*CEF*

(b) Use the average weight as a response. What factors appear to be influential?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	A	137.833	37996.1	12.0947
Error	B	-8.83333	156.056	0.049675
Error	C	11.6667	272.222	0.0866527
Model	D	-259.667	134854	42.926
Model	E	99.8333	19933.4	6.34511
Model	F	243.5	118585	37.7473
Error	AF	-34.3333	2357.56	0.750447
	Lenth's ME	563.698		

	Lenth's SME	1349.04



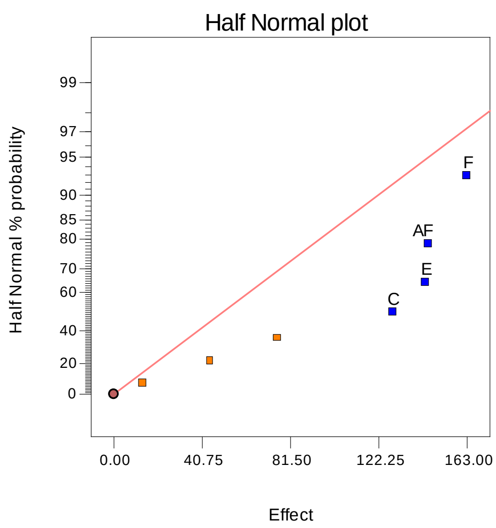
Factors *A*, *D*, *E* and *F* (and their aliases) are apparently important.

(c) Use the range of the weights as a response. What factors appear to be influential?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	44.5	3960.5	2.13311
Error	B	13.5	364.5	0.196319
Model	C	-129	33282	17.9256
Error	D	75.5	11400.5	6.14028
Model	E	144	41472	22.3367
Model	F	163	53138	28.62
Model	AF	145	42050	22.648
	Lenth's ME	728.384		
	Lenth's SME	1743.17		

Factors *C*, *E*, *F* and the *AF* interaction (and their aliases) appear to be large.



(d) What recommendations would you make to the process engineers?

It is not known exactly what to do here, since *A*, *D*, *E* and *F* are large effects, and because the design is resolution III, the main effects are aliased with two-factor interactions. Note, for example, that *D* is aliased with *EF* and the main effect could really be a *EF* interaction. If the main effects are really important, then setting all factors at the low level would minimize the amount of material stuck to the anodes. It would be necessary to run additional experiments to confirm these findings.

8.28. A 16-run experiment was performed in a semiconductor manufacturing plant to study the effects of six factors on the curvature or camber of the substrate devices produced. The six variables and their levels are shown in Table P8.2.

Table P8.2

Run	Lamination Temperature (c)	Lamination Time (s)	Lamination Pressure (tn)	Firing Temperature (c)	Firing Cycle Time (h)	Firing Dew Point (c)
1	55	10	5	1580	17.5	20
2	75	10	5	1580	29	26
3	55	25	5	1580	29	20
4	75	25	5	1580	17.5	26
5	55	10	10	1580	29	26
6	75	10	10	1580	17.5	20
7	55	25	10	1580	17.5	26
8	75	25	10	1580	29	20
9	55	10	5	1620	17.5	26
10	75	10	5	1620	29	20
11	55	25	5	1620	29	26
12	75	25	5	1620	17.5	20
13	55	10	10	1620	29	20
14	75	10	10	1620	17.5	26
15	55	25	10	1620	17.5	20
16	75	25	10	1620	29	26

Each run was replicated four times, and a camber measurement was taken on the substrate. The data are shown in Table P8.3.

Table P8.3

	Camber	for	Replicate	(in/in)	Total	Mean	Standard
Run	1	2	3	4	(10^{-4} in/in)	(10^{-4} in/in)	Deviation
1	0.0167	0.0128	0.0149	0.0185	629	157.25	24.418
2	0.0062	0.0066	0.0044	0.0020	192	48.00	20.976
3	0.0041	0.0043	0.0042	0.0050	176	44.00	4.083
4	0.0073	0.0081	0.0039	0.0030	223	55.75	25.025
5	0.0047	0.0047	0.0040	0.0089	223	55.75	22.410
6	0.0219	0.0258	0.0147	0.0296	920	230.00	63.639
7	0.0121	0.0090	0.0092	0.0086	389	97.25	16.029
8	0.0255	0.0250	0.0226	0.0169	900	225.00	39.420
9	0.0032	0.0023	0.0077	0.0069	201	50.25	26.725
10	0.0078	0.0158	0.0060	0.0045	341	85.25	50.341
11	0.0043	0.0027	0.0028	0.0028	126	31.50	7.681
12	0.0186	0.0137	0.0158	0.0159	640	160.00	20.083
13	0.0110	0.0086	0.0101	0.0158	455	113.75	31.120
14	0.0065	0.0109	0.0126	0.0071	371	92.75	29.510
15	0.0155	0.0158	0.0145	0.0145	603	150.75	6.750
16	0.0093	0.0124	0.0110	0.0133	460	115.00	17.450

(a) What type of design did the experimenters use?

The 2_{IV}^{6-2}, a 16-run design.

(b) What are the alias relationships in this design?

The defining relation is I=*ABCE*=*ACDF*=*BDEF*. The aliases are shown below.

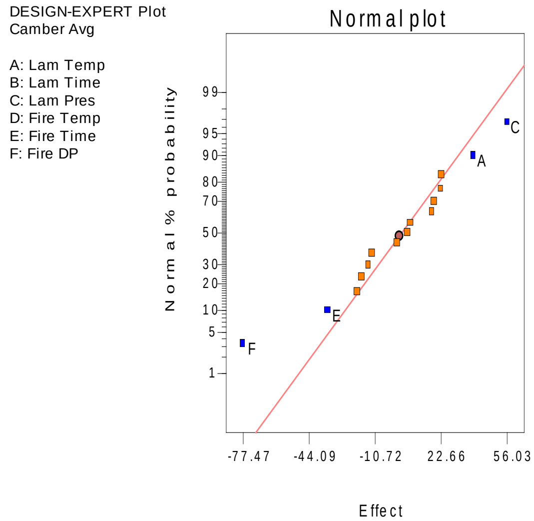
A(ABCE)=	*BCE*	*A(ACDF)=*	*CDF*	*A(BDEF)=*	*ABDEF*	*A=BCE=CDF=ABDEF*
B(ABCE)=	*ACE*	*B(ACDF)=*	*ABCDF*	*B(BDEF)=*	*DEF*	*B=ACE=ABCDF=DEF*
C(ABCE)=	*ABE*	*C(ACDF)=*	*ADF*	*C(BDEF)=*	*BCDEF*	*C=ABE=ADF=BCDEF*
D(ABCE)=	*ABCDE*	*D(ACDF)=*	*ACF*	*D(BDEF)=*	*BEF*	*D=ABCDE=ACF=BEF*
E(ABCE)=	*ABC*	*E(ACDF)=*	*ACDEF*	*E(BDEF)=*	*BDF*	*E=ABC=ACDEF=BDF*
F(ABCE)=	*ABCEF*	*F(ACDF)=*	*ACD*	*F(BDEF)=*	*BDE*	*F=ABCEF=ACD=BDE*
AB(ABCE)=	*CE*	*AB(ACDF)=*	*BCDF*	*AB(BDEF)=*	*ADEF*	*AB=CE=BCDF=ADEF*
AC(ABCE)=	*BE*	*AC(ACDF)=*	*DF*	*AC(BDEF)=*	*ABCDEF*	*AC=BE=DF=ABCDEF*
AD(ABCE)=	*BCDE*	*AD(ACDF)=*	*CF*	*AD(BDEF)=*	*ABEF*	*AD=BCDE=CF=ABEF*
AE(ABCE)=	*BC*	*AE(ACDF)=*	*CDEF*	*AE(BDEF)=*	*ABDF*	*AE=BC=CDEF=ABDF*
AF(ABCE)=	*BCEF*	*AF(ACDF)=*	*CD*	*AF(BDEF)=*	*ABDE*	*AF=BCEF=CD=ABDE*
BD(ABCE)=	*ACDE*	*BD(ACDF)=*	*ABCF*	*BD(BDEF)=*	*EF*	*BD=ACDE=ABCF=EF*
BF(ABCE)=	*ACEF*	*BF(ACDF)=*	*ABCD*	*BF(BDEF)=*	*DE*	*BF=ACEF=ABCD=DE*

(c) Do any of the process variables affect average camber?

Yes, per the analysis below, variables *A*, *C*, *E*, and *F* affect average camber.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	38.9063	6054.79	10.2962
Error	B	5.78125	133.691	0.227344
Model	C	56.0313	12558	21.355
Error	D	-14.2188	808.691	1.37519
Model	E	-34.4687	4752.38	8.08148
Model	F	-77.4688	24005.6	40.8219
Error	AB	19.1563	1467.85	2.49609
Error	AC	22.4063	2008.16	3.4149
Error	AD	-12.2188	597.191	1.01553
Error	AE	18.1563	1318.6	2.24229
Error	AF	-19.7187	1555.32	2.64483
Error	BC	Aliased		
Error	BD	23.0313	2121.75	3.60807
Error	BE	Aliased		
Error	BF	7.40625	219.41	0.37311
Error	CD	Aliased		
Error	CE	Aliased		
Error	CF	Aliased		
Error	DE	Aliased		
Error	DF	Aliased		
Error	EF	Aliased		
Error	ABC	Aliased		
Error	ABD	0.53125	1.12891	0.00191972
Error	ABE	Aliased		
Error	ABF	-17.3438	1203.22	2.04609
	Lenth's ME	71.9361		
	Lenth's SME	146.041		



Design Expert Output

Response: Camber Avg in in/in

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	47370.80	4	11842.70	11.39	0.0007	significant
A	*6054.79*	*1*	*6054.79*	*5.82*	*0.0344*	
C	*12558.00*	*1*	*12558.00*	*12.08*	*0.0052*	
E	*4752.38*	*1*	*4752.38*	*4.57*	*0.0558*	
F	*24005.63*	*1*	*24005.63*	*23.09*	*0.0005*	
Residual	11435.01	11	1039.55			
Cor Total	58805.81	15				

The Model F-value of 11.39 implies the model is significant. There is only a 0.07% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	32.24	R-Squared	0.8055
Mean	107.02	Adj R-Squared	0.7348
C.V.	30.13	Pred R-Squared	0.5886
PRESS	24193.08	Adeq Precision	11.478

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	107.02	1	8.06	89.27	124.76	
A-Lam Temp	19.45	1	8.06	1.71	37.19	1.00
C-Lam Pres	28.02	1	8.06	10.27	45.76	1.00
E-Fire Time	-17.23	1	8.06	-34.98	0.51	1.00
F-Fire DP	-38.73	1	8.06	-56.48	-20.99	1.00

Final Equation in Terms of Coded Factors:

Camber Avg =
+107.02
+19.45 * A
+28.02 * C
-17.23 * E
-38.73 * F

Final Equation in Terms of Actual Factors:

Camber Avg =
+263.17380
+1.94531 * Lam Temp
+11.20625 * Lam Pres

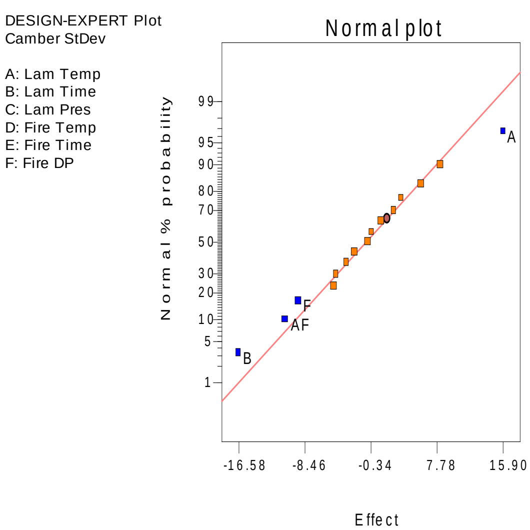
-2.99728	* Fire Time
-12.91146	* Fire DP

(d) Do any of the process variables affect the variability in camber measurements?

Yes, *A*, *B*, *F*, and *AF* interaction affect the variability in camber measurements.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	15.9035	1011.69	27.6623
Model	B	-16.5773	1099.22	30.0558
Error	C	5.8745	138.039	3.77437
Error	D	-3.2925	43.3622	1.18564
Error	E	-2.33725	21.851	0.597466
Model	F	-9.256	342.694	9.37021
Error	AB	0.95525	3.65001	0.0998014
Error	AC	2.524	25.4823	0.696757
Error	AD	-4.6265	85.618	2.34103
Error	AE	-0.18025	0.12996	0.00355347
Model	AF	-10.8745	473.019	12.9337
Error	BC	Aliased		
Error	BD	-4.85575	94.3132	2.57879
Error	BE	Aliased		
Error	BF	8.21825	270.159	7.38689
Error	CD	Aliased		
Error	CE	Aliased		
Error	CF	Aliased		
Error	DE	Aliased		
Error	DF	Aliased		
Error	EF	Aliased		
Error	ABC	Aliased		
Error	ABD	-0.68125	1.85641	0.0507593
Error	ABE	Aliased		
Error	ABF	3.39825	46.1924	1.26303
	Lenth's ME	17.8392		
	Lenth's SME	36.2162		



Response: Camber 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	2926.62	4	731.65	11.02	0.0008	significant

A	*1011.69*	*1*	*1011.69*	*15.23*	*0.0025*
B	*1099.22*	*1*	*1099.22*	*16.55*	*0.0019*
F	*342.69*	*1*	*342.69*	*5.16*	*0.0442*
AF	*473.02*	*1*	*473.02*	*7.12*	*0.0218*
Residual	730.65	11	66.42		
Cor Total	3657.27	15			

The Model F-value of 11.02 implies the model is significant. There is only a 0.08% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	8.15	R-Squared	0.8002
Mean	25.35	Adj R-Squared	0.7276
C.V.	32.15	Pred R-Squared	0.5773
PRESS	1545.84	Adeq Precision	9.516

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	25.35	1	2.04	20.87	29.84	
A-Lam Temp	7.95	1	2.04	3.47	12.44	1.00
B-Lam Time	-8.29	1	2.04	-12.77	-3.80	1.00
F-Fire DP	-4.63	1	2.04	-9.11	-0.14	1.00
AF	-5.44	1	2.04	-9.92	-0.95	1.00

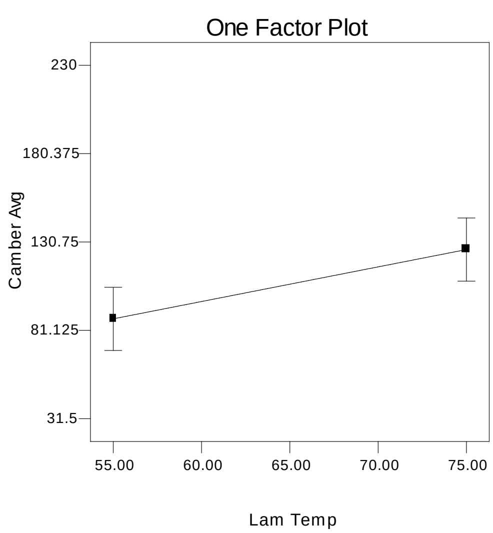
Final Equation in Terms of Coded Factors:

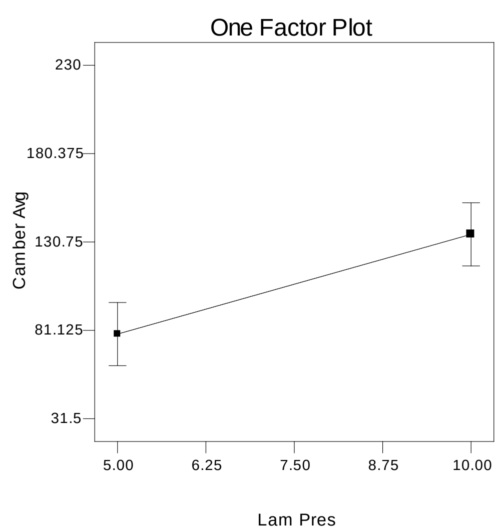
Camber StDev =
+25.35
+7.95 * A
-8.29 * B
-4.63 * F
-5.44 * A * F

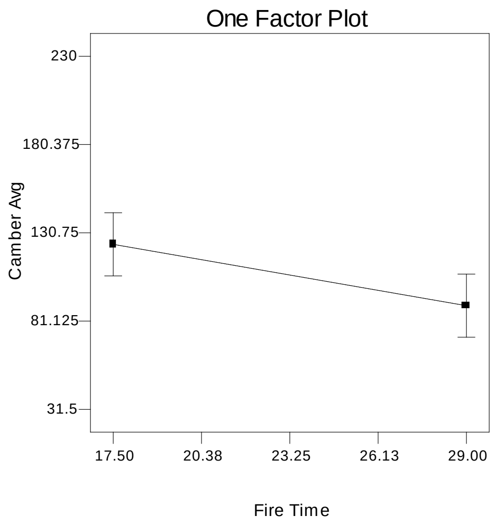
Final Equation in Terms of Actual Factors:

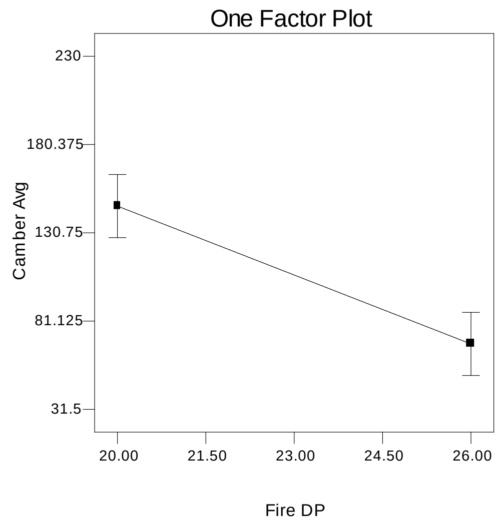
Camber StDev =
-242.46746
+4.96373 * Lam Temp
-1.10515 * Lam Time
+10.23804 * Fire DP
-0.18124 * Lam Temp * Fire DP

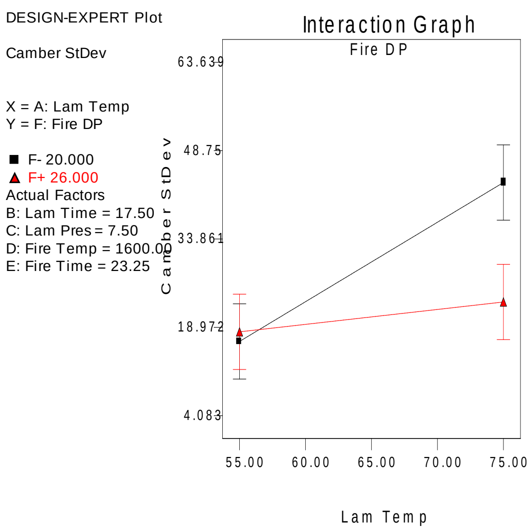
(e) If it is important to reduce camber as much as possible, what recommendations would you make?

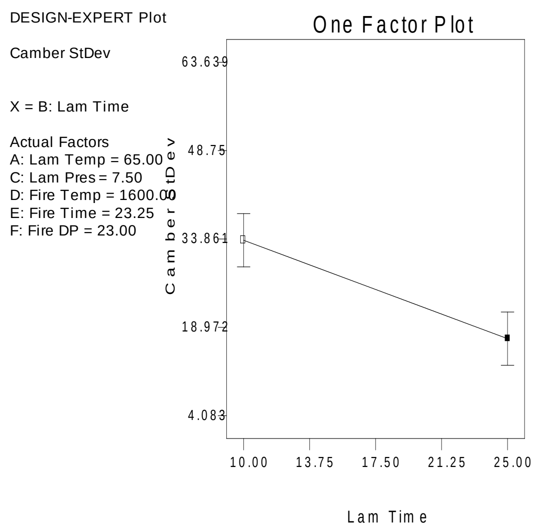












Run *A* and *C* at the low level and E and *F* at the high level. *B* at the low level enables a lower variation without affecting the average camber.

8.29. A spin coater is used to apply photoresist to a bare silicon wafer. This operation usually occurs early in the semiconductor manufacturing process, and the average coating thickness and the variability in the coating thickness has an important impact on downstream manufacturing steps. Six variables are used in the experiment. The variables and their high and low levels are as follows:

Factor	Low Level	High Level
Final Spin Speed	7350 rpm	6650 rpm
Acceleration Rate	5	20
Volume of Resist Applied	3 cc	5 cc
Time of Spin	14 s	6 s
Resist Batch Variation	Batch 1	Batch 2
Exhaust Pressure	Cover Off	Cover On

The experimenter decides to use a 2^{6-1} design and to make three readings on resist thickness on each test wafer. The data are shown in table P8.4.

Table P8.4

	A	*B*	*C*	*D*	*E*	*F*	Resist Thickness				
Run	Volume	Batch	Time	Speed	Acc.	Cover	Left	Center	Right	Avg.	Range
1	5	2	14	7350	5	Off	4531	4531	4515	4525.7	16
2	5	1	6	7350	5	Off	4446	4464	4428	4446	36
3	3	1	6	6650	5	Off	4452	4490	4452	4464.7	38
4	3	2	14	7350	20	Off	4316	4328	4308	4317.3	20
5	3	1	14	7350	5	Off	4307	4295	4289	4297	18
6	5	1	6	6650	20	Off	4470	4492	4495	4485.7	25
7	3	1	6	7350	5	On	4496	4502	4482	4493.3	20
8	5	2	14	6650	20	Off	4542	4547	4538	4542.3	9
9	5	1	14	6650	5	Off	4621	4643	4613	4625.7	30
10	3	1	14	6650	5	On	4653	4670	4645	4656	25
11	3	2	14	6650	20	On	4480	4486	4470	4478.7	16
12	3	1	6	7350	20	Off	4221	4233	4217	4223.7	16
13	5	1	6	6650	5	On	4620	4641	4619	4626.7	22
14	3	1	6	6650	20	On	4455	4480	4466	4467	25
15	5	2	14	7350	20	On	4255	4288	4243	4262	45
16	5	2	6	7350	5	On	4490	4534	4523	4515.7	44
17	3	2	14	7350	5	On	4514	4551	4540	4535	37
18	3	1	14	6650	20	Off	4494	4503	4496	4497.7	9
19	5	2	6	7350	20	Off	4293	4306	4302	4300.3	13
20	3	2	6	7350	5	Off	4534	4545	4512	4530.3	33
21	5	1	14	6650	20	On	4460	4457	4436	4451	24
22	3	2	6	6650	5	On	4650	4688	4656	4664.7	38
23	5	1	14	7350	20	Off	4231	4244	4230	4235	14
24	3	2	6	7350	20	On	4225	4228	4208	4220.3	20
25	5	1	14	7350	5	On	4381	4391	4376	4382.7	15
26	3	2	6	6650	20	Off	4533	4521	4511	4521.7	22
27	3	1	14	7350	20	On	4194	4230	4172	4198.7	58
28	5	2	6	6650	5	Off	4666	4695	4672	4677.7	29
29	5	1	6	7350	20	On	4180	4213	4197	4196.7	33
30	5	2	6	6650	20	On	4465	4496	4463	4474.7	33
31	5	2	14	6650	5	On	4653	4685	4665	4667.7	32
32	3	2	14	6650	5	Off	4683	4712	4677	4690.7	35

(a) Verify that this is a 2^{6-1} design. Discuss the alias relationships in this design.

I=*ABCDEF*. This is a resolution VI design where main effects are aliased with five-factor interactions and two-factor interactions are aliased with four-factor interactions.

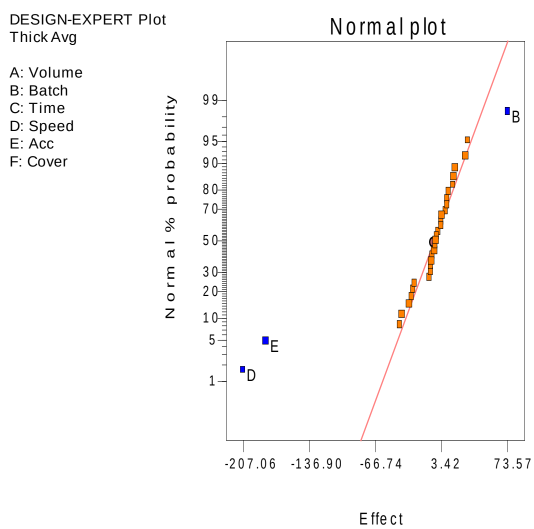
(b) What factors appear to affect average resist thickness?

Factors *B*, *D*, and *E* appear to affect the average resist thickness.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	9.925	788.045	0.107795
Model	B	73.575	43306.2	5.92378
Error	C	3.375	91.125	0.0124648
Model	D	-207.062	342999	46.9182
Model	E	-182.925	267692	36.6172
Error	F	-5.6625	256.511	0.0350877
Error	AB	-9	648	0.0886387
Error	AC	-7.3	426.32	0.0583155
Error	AD	-3.8625	119.351	0.0163258
Error	AE	-7.1	403.28	0.0551639
Error	AF	-26.9875	5826.6	0.79701
Error	BC	10.875	946.125	0.129419
Error	BD	18.1125	2624.5	0.359001
Error	BE	-28.35	6429.78	0.879518
Error	BF	-30.2375	7314.45	1.00053

Error	CD	-24.9875	4995	0.683257
Error	CE	8.2	537.92	0.0735811
Error	CF	-6.7875	368.561	0.0504148
Error	DE	-38.5375	11881.1	1.6252
Error	DF	-3.2	81.92	0.0112057
Error	EF	-41.1625	13554.8	1.85414
Error	ABC	0.375	1.125	0.000153887
Error	ABD	Aliased		
Error	ABE	16.5	2178	0.297925
Error	ABF	31.4125	7893.96	1.0798
Error	ACD	15.5875	1943.76	0.265883
Error	ACE	Aliased		
Error	ACF	Aliased		
Error	ADE	9.5375	727.711	0.0995423
Error	ADF	Aliased		
Error	AEF	Aliased		
Error	BCD	29.0875	6768.66	0.925873
Error	BCE	-1.625	21.125	0.00288965
Error	BCF	Aliased		
Error	BDE	-1.8875	28.5013	0.00389863
Error	BDF	3.95	124.82	0.0170739
Error	BEF	Aliased		
Error	CDE	Aliased		
Error	CDF	Aliased		
Error	CEF	3.1375	78.7512	0.0107722
Error	DEF	Aliased		
	Lenth's ME	28.6178		
	Lenth's SME	54.4118		



Design Expert Output

Response: Thick Avg
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	6.540E+005	3	2.180E+005	79.21	< 0.0001	significant
B	*43306.24*	*1*	*43306.24*	*15.74*	*0.0005*	
D	*3.430E+005*	*1*	*3.430E+005*	*124.63*	*< 0.0001*	
E	*2.677E+005*	*1*	*2.677E+005*	*97.27*	*< 0.0001*	
Residual	77059.83	28	2752.14			
Cor Total	7.311E+005	31				

The Model F-value of 79.21 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	52.46	R-Squared	0.8946
Mean	4458.51	Adj R-Squared	0.8833
C.V.	1.18	Pred R-Squared	0.8623
PRESS	1.006E+005	Adeq Precision	24.993

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	4458.51	1	9.27	4439.52	4477.51	
B-Batch	36.79	1	9.27	17.79	55.78	1.00
D-Speed	-103.53	1	9.27	-122.53	-84.53	1.00
E-Acc	-91.46	1	9.27	-110.46	-72.47	1.00

Final Equation in Terms of Coded Factors:

Thick Avg =
+4458.51
+36.79 * B
-103.53 * D
-91.46 * E

Final Equation in Terms of Actual Factors:

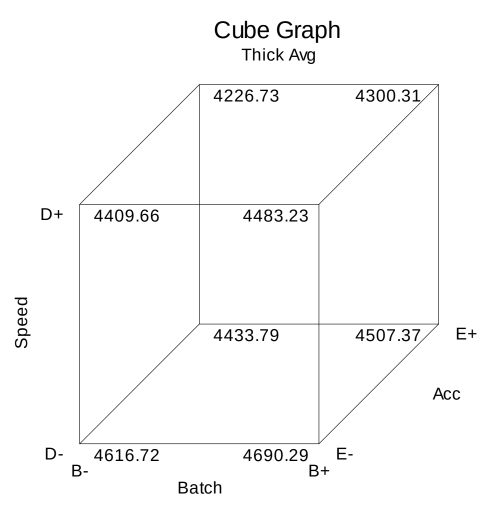
Batch Batch 1
Thick Avg =
+6644.78750
-0.29580 * Speed
-12.19500 * Acc

Batch Batch 2
Thick Avg =
+6718.36250
-0.29580 * Speed
-12.19500 * Acc

(c) Since the volume of resist applied has little effect on average thickness, does this have any important practical implications for the process engineers?

Yes, less material could be used.

(d) Project this design into a smaller design involving only the significant factors. Graphically display the results. Does this aid in interpretation?

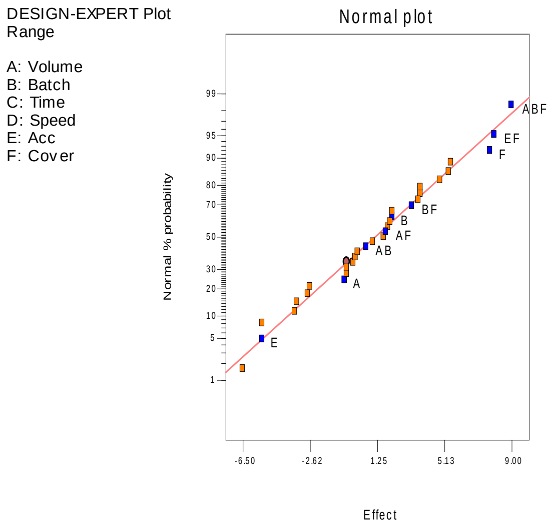


The cube plot usually assists the experimenter in drawing conclusions.

(e) Use the range of resist thickness as a response variable. Is there any indication that any of these factors affect the variability in resist thickness?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	-0.625	3.125	0.0777387
Model	B	2.125	36.125	0.89866
Error	C	-2.75	60.5	1.50502
Error	D	1.625	21.125	0.525514
Model	E	-5.375	231.125	5.74956
Model	F	7.75	480.5	11.9531
Model	AB	0.625	3.125	0.0777387
Error	AC	-3.5	98	2.43789
Error	AD	-0.125	0.125	0.00310955
Error	AE	1.875	28.125	0.699649
Model	AF	1.75	24.5	0.609472
Error	BC	0	0	0
Error	BD	0.125	0.125	0.00310955
Error	BE	-5.375	231.125	5.74956
Model	BF	3.25	84.5	2.10206
Error	CD	3.75	112.5	2.79859
Error	CE	3.75	112.5	2.79859
Error	CF	4.875	190.125	4.72962
Error	DE	5.375	231.125	5.74956
Error	DF	5.5	242	6.02009
Model	EF	8	512	12.7367
Error	ABC	Aliased		
Error	ABD	Aliased		
Error	ABE	3.625	105.125	2.61513
Model	ABF	9	648	16.1199
Error	ACD	-6.5	338	8.40822
Error	ACE	Aliased		
Error	ACF	Aliased		
Error	ADE	-3.375	91.125	2.26686
Error	ADF	-0.5	2	0.0497528
Error	AEF	1	8	0.199011
Error	BCD	Aliased		
Error	BCE	Aliased		
Error	BCF	Aliased		
Error	BDE	-2.625	55.125	1.37131
Error	BDF	-0.5	2	0.0497528
Error	BEF	Aliased		
Error	CDE	Aliased		
Error	CDF	Aliased		
Error	CEF	2.125	36.125	0.89866
Error	DEF	2	32	0.796045
	Lenth's ME	9.15104		
	Lenth's SME	17.3991		



Design Expert Output

Response: Thick 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	2023.00	9	224.78	2.48	0.0400	significant
A	*3.13*	*1*	*3.13*	*0.034*	*0.8545*	
B	*36.13*	*1*	*36.13*	*0.40*	*0.5346*	
E	*231.12*	*1*	*231.12*	*2.55*	*0.1248*	
F	*480.50*	*1*	*480.50*	*5.29*	*0.0313*	
AB	*3.12*	*1*	*3.12*	*0.034*	*0.8545*	
AF	*24.50*	*1*	*24.50*	*0.27*	*0.6086*	
BF	*84.50*	*1*	*84.50*	*0.93*	*0.3451*	
EF	*512.00*	*1*	*512.00*	*5.64*	*0.0267*	
ABF	*648.00*	*1*	*648.00*	*7.14*	*0.0139*	
Residual	1996.88	22	90.77			
Cor Total	4019.88	31				

The Model F-value of 2.48 implies the model is significant. There is only a 4.00% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	9.53	R-Squared	0.5032
Mean	26.56	Adj R-Squared	0.3000
C.V.	35.87	Pred R-Squared	-0.0510
PRESS	4224.79	Adeq Precision	5.586

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	26.56	1	1.68	23.07	30.06	
A-Volume	-0.31	1	1.68	-3.81	3.18	1.00
B-Batch	1.06	1	1.68	-2.43	4.56	1.00
E-Acc	-2.69	1	1.68	-6.18	0.81	1.00
F-Cover	3.88	1	1.68	0.38	7.37	1.00
AB	0.31	1	1.68	-3.18	3.81	1.00
AF	0.88	1	1.68	-2.62	4.37	1.00
BF	1.63	1	1.68	-1.87	5.12	1.00
EF	4.00	1	1.68	0.51	7.49	1.00
ABF	4.50	1	1.68	1.01	7.99	1.00

Final Equation in Terms of Coded Factors:

Thick Range =
+26.56
-0.31 * A

```
              +1.06      * B
              -2.69      * E
              +3.88      * F
              +0.31      * A * B
              +0.88      * A * F
              +1.63      * B * F
              +4.00      * E * F
              +4.50      * A * B * F
```

Final Equation in Terms of Actual Factors:

```
              Batch    Batch 1
              Cover    Off
         Thick Range     =
           +22.39583
            +3.00000     * Volume
            -0.89167     * Acc

              Batch    Batch 2
              Cover    Off
         Thick Range     =
           +54.77083
            -5.37500     * Volume
            -0.89167     * Acc

              Batch    Batch 1
              Cover    On
         Thick Range     =
           +42.56250
            -4.25000     * Volume
            +0.17500     * Acc

              Batch    Batch 2
              Cover    On
         Thick Range     =
            +9.43750
            +5.37500     * Volume
            +0.17500     * Acc
```

The model for thickness range is not very strong. Notice the small value of R^2, and in particular, the adjusted R^2. Often we find that obtaining a good model for a response that expresses variability is not as easy as finding a satisfactory model for a response that measures the mean.

(f) Where would you recommend that the process engineers run the process?

Considering only the average thickness results, the engineers could use factors *B*, *D* and *E* to put the process mean at target. Then the engineer could consider the other factors on the range model to try to set the factors to reduce the variation in thickness at that mean.

8.30. Harry Peterson-Nedry (a friend of the author) owns a vineyard and winery in Newberg, Oregon. They grow several varieties of grapes and manufacture wine. Harry and Judy have used factorial designs for process and product development in the winemaking segment of their business. This problem describes the experiment conducted for their 1985 Pinot Noir. Eight variables, shown in Table P8.5, were originally studied in this experiment:

Table P8.5

	Variable	Low Level	High Level
A	Pinot Noir Clone	Pommard	Wadenswil
B	Oak Type	Allier	Troncais
C	Age of Barrel	Old	New
D	Yeast/Skin Contact	Champagne	Montrachet
E	Stems	None	All

F	Barrel Toast	Light	Medium
G	Whole Cluster	None	10%
H	Fermentation Temperature	Low (75 F Max)	High (92 F Max)

Harry decided to use a 2^{8-4}_{IV} design with 16 runs. The wine was taste-tested by a panel of experts on 8 March 1986. Each expert ranked the 16 samples of wine tasted, with rank 1 being the best. The design and taste-test panel results are shown in Table P8.6.

Table P8.6

Run	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	HPN	JPN	CAL	DCM	RGB	y_{bar}	*s*
1	-	-	-	-	-	-	-	-	12	6	13	10	7	9.6	3.05
2	+	-	-	-	-	+	+	+	10	7	14	14	9	10.8	3.11
3	-	+	-	-	+	-	+	+	14	13	10	11	15	12.6	2.07
4	+	+	-	-	+	+	-	-	9	9	7	9	12	9.2	1.79
5	-	-	+	-	+	+	+	-	8	8	11	8	10	9.0	1.41
6	+	-	+	-	+	-	-	+	16	12	15	16	16	15.0	1.73
7	-	+	+	-	-	+	-	+	6	5	6	5	3	5.0	1.22
8	+	+	+	-	-	-	+	-	15	16	16	15	14	15.2	0.84
9	-	-	-	+	+	+	-	+	1	2	3	3	2	2.2	0.84
10	+	-	-	+	+	-	+	-	7	11	4	7	6	7.0	2.55
11	-	+	-	+	-	+	+	-	13	3	8	12	8	8.8	3.96
12	+	+	-	+	-	-	-	+	3	1	5	1	4	2.8	1.79
13	-	-	+	+	-	-	+	+	2	10	2	4	5	4.6	3.29
14	+	-	+	+	-	+	-	-	4	4	1	2	1	2.4	1.52
15	-	+	+	+	+	-	-	-	5	15	9	6	11	9.2	4.02
16	+	+	+	+	+	+	+	+	11	14	12	13	13	12.6	1.14

(a) What are the alias relationships in the design selected by Harry?

$$E = BCD,\ F = ACD,\ G = ABC,\ H = ABD$$

$$\text{Defining Contrast}: \quad I = BCDE = ACDF = ABEF = ABCG = ADEG = BDFG = CEFG = ABDH$$
$$= ACEH = BCFH = DEFH = CDGH = BEGH = AFGH = ABCDEFGH$$

Aliases:

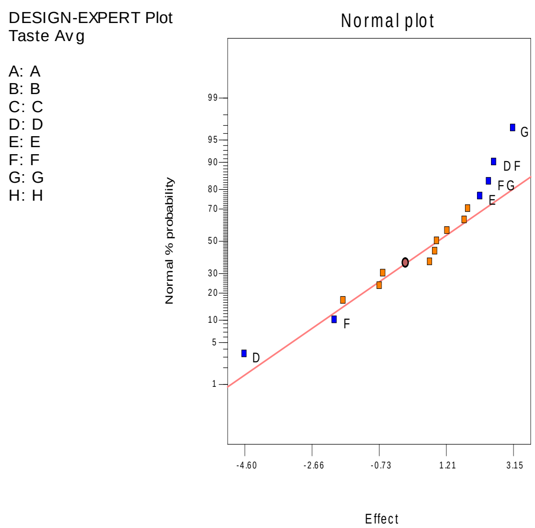
$$A = BCG = BDH = BEF = CDF = CEH = DEG = FGH$$
$$B = ACG = ADH = AEF = CDE = CFH = DFG = EGH$$
$$C = ABG = ADF = AEH = BDE = BFH = DGH = EFG$$
$$D = ABH = ACF = AEG = BCE = BFG = CGH = EFH$$
$$E = ABF = ACH = ADG = BCD = BGH = CFG = DFH$$
$$F = ABE = ACD = AGH = BCH = BDG = CEG = DEH$$
$$G = ABC = ADE = AFH = BDF = BEH = CDH = CEF$$
$$H = ABD = ACE = AFG = BCF = BEG = CDG = DEF$$
$$AB = CG = DH = EF$$
$$AC = BG = DF = EH$$
$$AD = BH = CF = EG$$
$$AE = BF = CH = DG$$
$$AF = BE = CD = GH$$
$$AG = BC = DE = FH$$
$$AH = BD = CE = FG$$

(b) Use the average ranks ($\bar{y}$) as a response variable. Analyze the data and draw conclusions. You will find it helpful to examine a normal probability plot of effect estimates.

The effects list and normal probability plot of effects are shown below. Factors *D*, *E*, *F*, and *G* appear to be significant. Also note that the *DF* and *FG* interactions were chosen instead of *AC* and *AH* based on the alias structure shown above.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Require	Intercept			
Error	A	1.75	12.25	4.57636
Error	B	1.85	13.69	5.11432
Error	C	1.25	6.25	2.33488
Model	D	-4.6	84.64	31.6198
Model	E	2.2	19.36	7.23252
Model	F	-2	16	5.97729
Model	G	3.15	39.69	14.8274
Error	H	-0.6	1.44	0.537956
Error	AB	-0.7	1.96	0.732218
Ignore	AC	Aliased		
Error	AD	-1.75	12.25	4.57636
Error	AE	0.95	3.61	1.34863
Error	AF	0.75	2.25	0.840556
Error	AG	0.9	3.24	1.2104
Ignore	AH	Aliased		
Model	DF	2.6	27.04	10.1016
Model	FG	2.45	24.01	8.96967
	Lenth's ME	6.74778		
	Lenth's SME	13.699		



Design Expert Output

Response: Taste Avg
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	210.74	6	35.12	5.55	0.0115	significant
D	*84.64*	*1*	*84.64*	*13.38*	*0.0053*	
E	*19.36*	*1*	*19.36*	*3.06*	*0.1142*	
F	*16.00*	*1*	*16.00*	*2.53*	*0.1462*	
G	*39.69*	*1*	*39.69*	*6.27*	*0.0336*	
DF	*27.04*	*1*	*27.04*	*4.27*	*0.0687*	
FG	*24.01*	*1*	*24.01*	*3.80*	*0.0832*	
Residual	56.94	9	6.33			
Cor Total	267.68	15				

The Model F-value of 5.55 implies the model is significant. There is only a 1.15% chance that a "Model F-Value" this large could occur due to noise.

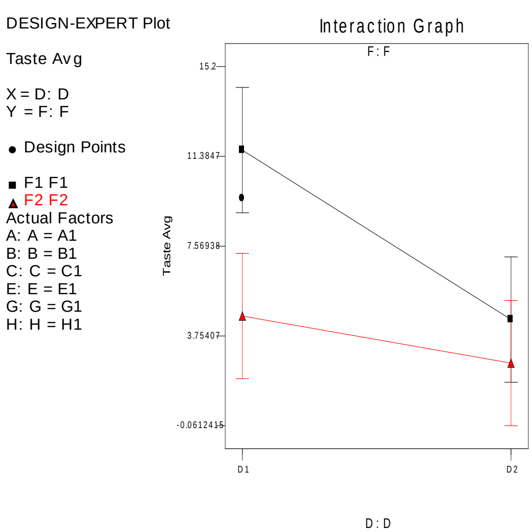
Std. Dev.	2.52	R-Squared	0.7873
Mean	8.50	Adj R-Squared	0.6455
C.V.	29.59	Pred R-Squared	0.3277
PRESS	179.96	Adeq Precision	7.183

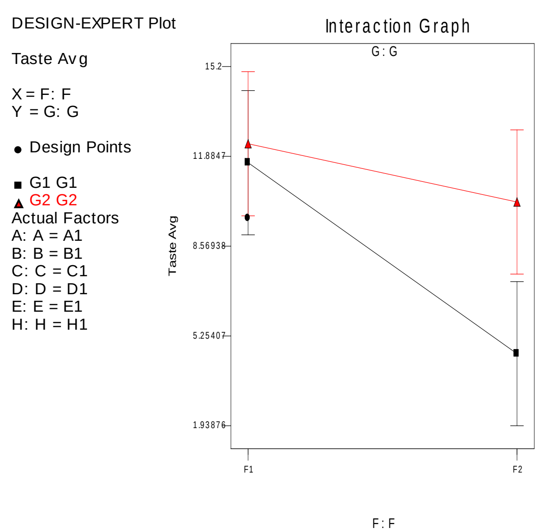
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	8.50	1	0.63	7.08	9.92	
D-D	-2.30	1	0.63	-3.72	-0.88	1.00
E-E	1.10	1	0.63	-0.32	2.52	1.00
F-F	-1.00	1	0.63	-2.42	0.42	1.00
G-G	1.57	1	0.63	0.15	3.00	1.00
DF	1.30	1	0.63	-0.12	2.72	1.00
FG	1.23	1	0.63	-0.20	2.65	1.00

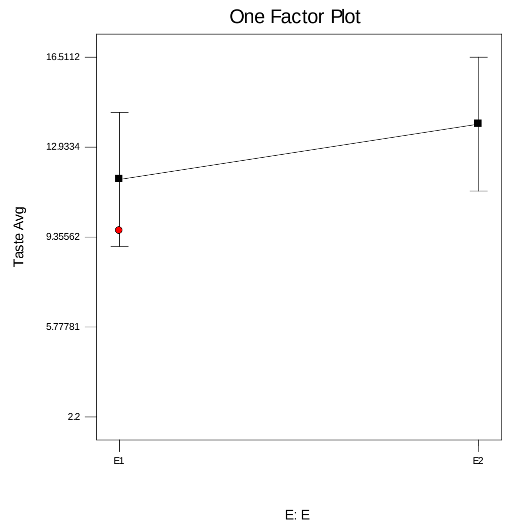
Final Equation in Terms of Coded Factors:

Taste Avg =
+8.50
-2.30 * D
+1.10 * E
-1.00 * F
+1.57 * G
+1.30 * D * F
+1.23 * F * G

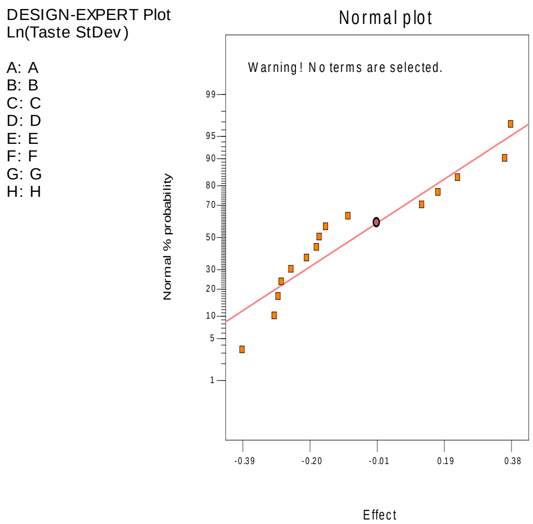
Factors *D* and *G* and are important. Factor *E* and the *DF* and *FG* interactions are moderately important and were included in the model because the PRESS statistic showed improvement with their inclusion. Factor *F* is added to the model to preserve hierarchy. As stated earlier, the interactions are aliased with other two-factor interactions that could also be important. So the interpretation of the two-factor interaction is somewhat uncertain. Normally, we would add runs to the design to isolate the significant interactions, but that will not work very well here because each experiment requires a full growing season. In other words, it would require a very long time to add runs to de-alias the alias chains of interest.







(c) Use the standard deviation of the ranks (or some appropriate transformation such as log *s*) as a response variable. What conclusions can you draw about the effects of the eight variables on variability in wine quality?



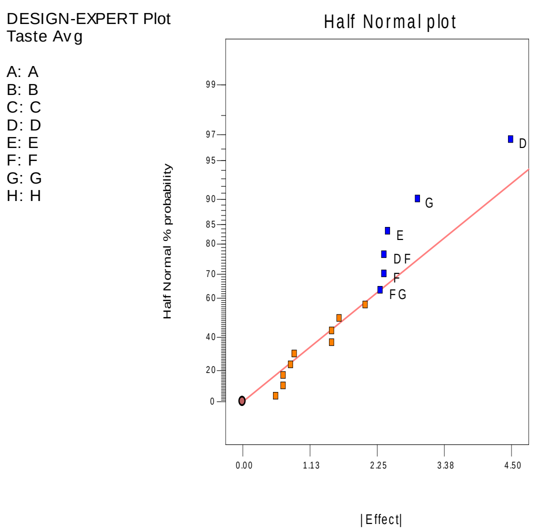
There do not appear to be any significant factors.

(d) After looking at the results, Harry decides that one of the panel members (DCM) knows more about beer than he does about wine, so they decide to delete his ranking. What affect would this have on the results and on conclusions from parts (b) and (c)?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Require	Intercept			
Error	A	1.625	10.5625	4.02957
Error	B	2.0625	17.0156	6.49142
Error	C	1.5	9	3.43348
Model	D	-4.5	81	30.9013
Model	E	2.4375	23.7656	9.06652

Model	F	-2.375	22.5625	8.60753
Model	G	2.9375	34.5156	13.1676
Error	H	-0.6875	1.89063	0.721268
Error	AB	-0.5625	1.26563	0.482833
Ignore	AC	Aliased		
Error	AD	-1.5	9	3.43348
Error	AE	0.6875	1.89063	0.721268
Error	AF	0.875	3.0625	1.16834
Error	AG	0.8125	2.64062	1.00739
Ignore	AH	Aliased		
Model	DF	2.375	22.5625	8.60753
Model	FG	2.3125	21.3906	8.16047
	Lenth's ME	6.26579		
	Lenth's SME	12.7205		



Design Expert Output

Response: Taste Avg

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	205.80	6	34.30	5.48	0.0120	significant
D	*81.00*	*1*	*81.00*	*12.94*	*0.0058*	
E	*23.77*	*1*	*23.77*	*3.80*	*0.0831*	
F	*22.56*	*1*	*22.56*	*3.60*	*0.0901*	
G	*34.52*	*1*	*34.52*	*5.51*	*0.0434*	
DF	*22.56*	*1*	*22.56*	*3.60*	*0.0901*	
FG	*21.39*	*1*	*21.39*	*3.42*	*0.0975*	
Residual	56.33	9	6.26			
Cor Total	262.13	15				

The Model F-value of 5.48 implies the model is significant. There is only a 1.20% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.50	R-Squared	0.7851
Mean	8.50	Adj R-Squared	0.6418
C.V.	29.43	Pred R-Squared	0.3208
PRESS	178.02	Adeq Precision	7.403

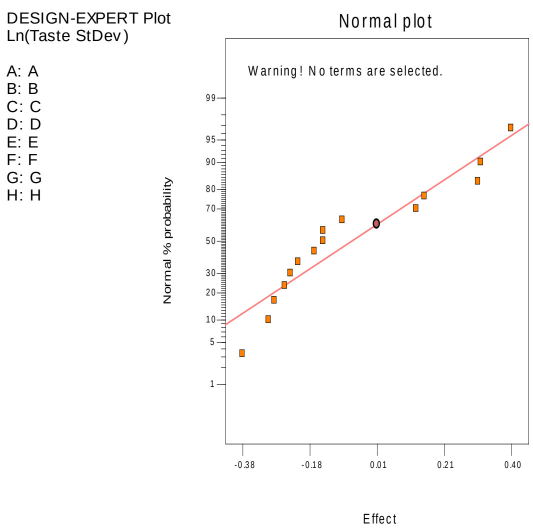
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	8.50	1	0.63	7.09	9.91	
D-D	-2.25	1	0.63	-3.66	-0.84	1.00
E-E	1.22	1	0.63	-0.20	2.63	1.00

F-F	-1.19	1	0.63	-2.60	0.23	1.00
G-G	1.47	1	0.63	0.054	2.88	1.00
DF	1.19	1	0.63	-0.23	2.60	1.00
FG	1.16	1	0.63	-0.26	2.57	1.00

Final Equation in Terms of Coded Factors:

Taste Avg	=
+8.50	
-2.25	* D
+1.22	* E
-1.19	* F
+1.47	* G
+1.19	* D * F
+1.16	* F * G

The results are very similar for average taste without DCM as they were with DCM.



The standard deviation response is much the same with or without DCM's responses. Again, there are no significant factors.

(e) Suppose that just before the start of the experiment, Harry discovered that the eight new barrels they ordered from France for use in the experiment would not arrive in time, and all 16 runs would have to be made with old barrels. If Harry just drops column C from their design, what does this do to the alias relationships? Do they need to start over and construct a new design?

The resulting design is a 2_{IV}^{7-3} with defining relations: $I = ABEF = ADEG = BDFG = ABDH = DEFH = BEGH = AFGH$.

(f) Harry knows from experience that some treatment combinations are unlikely to produce good results. For example, the run with all eight variables at the high level generally results in a poorly rated wine. This was confirmed in the 8 March 1986 taste test. He wants to set up a new design for their 1986 Pinot Noir using these same eight variables, but they do not want to make the run with all eight factors at the high level. What design would you suggest?

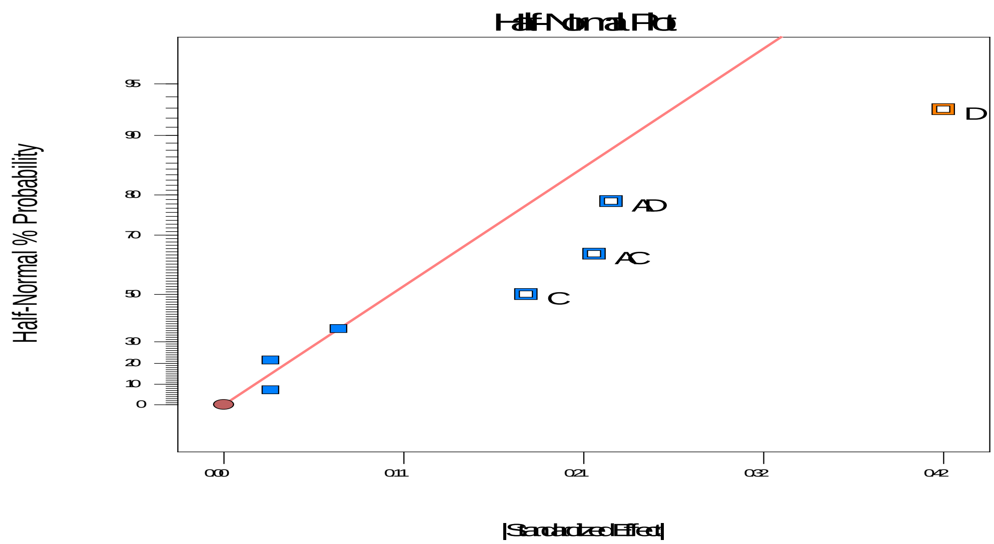
By changing the sign of any of the design generators, a design that does not include the principal fraction will be generated. This will give a design without an experimental run combination with all of the variables at the high level.

8.31. Consider the isatin yield data from the experiment described in Problem 6.38. The original experiment was a 2^4 full factorial. Suppose that the original experimenters could only afford eight runs. Set up the 2^{4-1} fractional factorial with $I = ABCD$ and select the responses for the runs from the full factorial data in Problem 6.38. Analyze and draw conclusions. Compare your findings with those from the full factorial in Problem 6.38.

The runs for the 2^{4-1} fractional factorial with $I = ABCD$ are shown below.

Acid strength	Reaction time	Amount of acid	Reaction temperature	Yield
-1	-1	-1	-1	6.08
1	-1	-1	1	6.68
-1	1	-1	1	6.73
1	1	-1	-1	6.43
-1	-1	1	1	6.77
1	-1	1	-1	6.09
-1	1	1	-1	6.12
1	1	1	1	6.23

Based on the analysis shown below, the *C* and *D* factors along with the *AC* and *AD* interactions are significant. The normal probability plot from Problem 6.38 identified factors *A*, *D*, the *AD* and *BD* interactions as being significant. In the 2^{4-1} fractional factorial with $I = ABCD$, the AC and *BD* interactions are aliased. Based on the analysis of the full factorial experiment, the *BD* interaction should be chosen below rather than the *AC*. However, if we assume that the experimenter was not aware of this, additional runs should be performed to de-alias the *AC* and *BD* interactions.



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	0.63	5	0.13	82.94	0.0120	significant
A-Acid strength	*9.113E-003*	*1*	*9.113E-003*	*6.02*	*0.1335*	
C-Amount of acid	*0.063*	*1*	*0.063*	*41.66*	*0.0232*	
D-Reaction temperature	*0.36*	*1*	*0.36*	*236.04*	*0.0042*	
AC	*0.095*	*1*	*0.095*	*62.55*	*0.0156*	

AD	*0.10*	*1*	*0.10*	*68.44*	*0.0143*
Residual	3.025E-003	2	1.512E-003		
Cor Total	0.63	7			

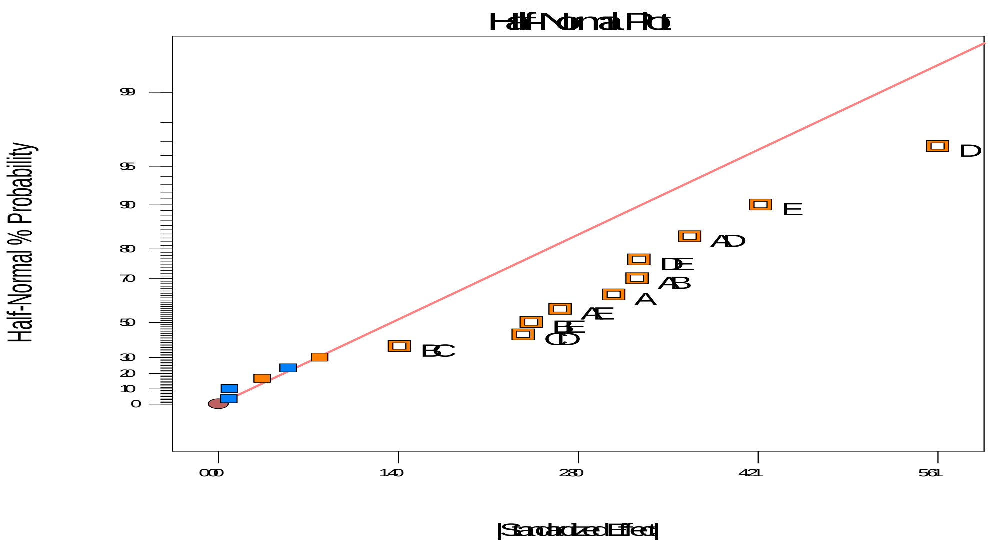
8.32. Consider the 2^5 factorial in Problem 6.39. Suppose that the experimenters could only afford 16 runs. Set up the 2^{5-1} fractional factorial design with $I=ABCDE$ and select the responses from the full factorial data in Problem 6.39.

The runs for the 2^{5-1} fractional factorial design with $I=ABCDE$ are shown below.

A	B	C	D	E	y
-1	-1	-1	-1	1	7.93
1	-1	-1	-1	-1	5.56
-1	1	-1	-1	-1	5.77
1	1	-1	-1	1	12
-1	-1	1	-1	-1	9.17
1	-1	1	-1	1	3.65
-1	1	1	-1	1	6.4
1	1	1	-1	-1	5.69
-1	-1	-1	1	-1	8.82
1	-1	-1	1	1	17.55
-1	1	-1	1	1	8.87
1	1	-1	1	-1	8.94
-1	-1	1	1	1	13.06
1	-1	1	1	-1	11.49
-1	1	1	1	-1	6.25
1	1	1	1	1	26.05

(a) Analyze the data and draw conclusions.

The analysis which identifies the significant effects is shown below.



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	460.75	12	38.40	92.45	0.0016	significant
A-A	*38.01*	*1*	*38.01*	*91.51*	*0.0024*	
B-B	*0.47*	*1*	*0.47*	*1.13*	*0.3658*	
C-C	*2.50*	*1*	*2.50*	*6.01*	*0.0915*	
D-D	*125.78*	*1*	*125.78*	*302.84*	*0.0004*	
E-E	*71.49*	*1*	*71.49*	*172.13*	*0.0010*	
AB	*42.64*	*1*	*42.64*	*102.67*	*0.0020*	
AD	*54.02*	*1*	*54.02*	*130.08*	*0.0014*	
AE	*28.41*	*1*	*28.41*	*68.40*	*0.0037*	
BC	*7.98*	*1*	*7.98*	*19.22*	*0.0220*	
BE	*23.81*	*1*	*23.81*	*57.34*	*0.0048*	
CD	*22.61*	*1*	*22.61*	*54.44*	*0.0051*	
DE	*43.03*	*1*	*43.03*	*103.62*	*0.0020*	
Residual	1.25	3	0.42			
Cor Total	461.99	15				

(b) Compare your findings with those from the full factorial in Problem 6.39.

The results are similar to those found in Problem 6.39 with the exception that the *BC* and *CD* interactions are now significant, and factor *C* remains in the model because of these interactions.

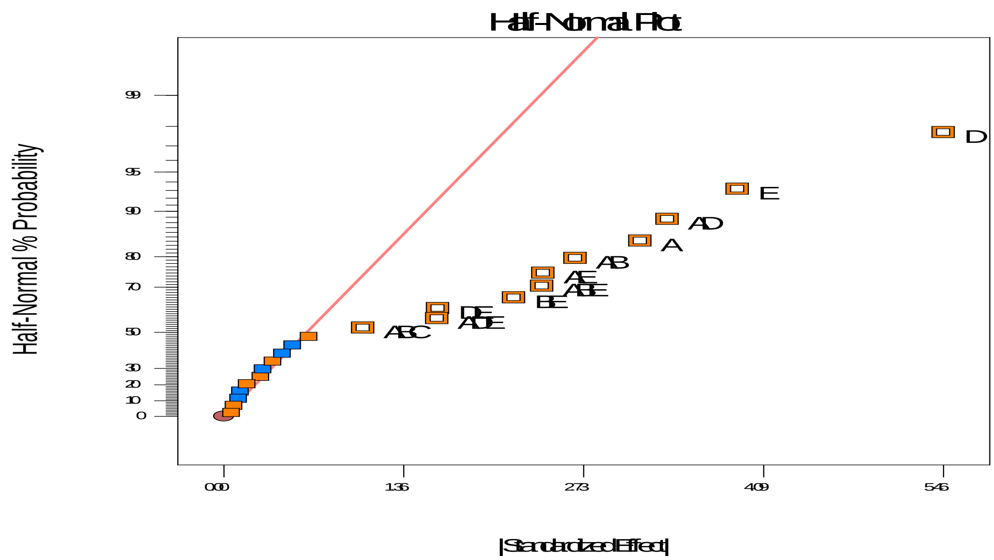
(c) Are there any potential interactions that need further study? What additional runs do you recommend? Select these runs from the full factorial design in Problem 6.39 and analyze the new design. Discuss your conclusions.

The *BC* interaction is aliased with the *ADE* interaction, and the *CD* interaction is aliased with the *ABE* interaction. The *ADE* and *ABE* interactions in Problem 6.39 were both found to be significant. Including these effects would also allow us to remove the insignificant factor *C* from the model.

The additional runs are shown below as block 2. This is a partial fold-over where factor A was chosen for the fold-over, and factor C at the -1 level was chosen for the runs. This allowed both the ABE and ADE interactions to be de-aliased.

Block	A	B	C	D	E	y
1	-1	-1	-1	-1	1	7.93
1	1	-1	-1	-1	-1	5.56
1	-1	1	-1	-1	-1	5.77
1	1	1	-1	-1	1	12
1	-1	-1	1	-1	-1	9.17
1	1	-1	1	-1	1	3.65
1	-1	1	1	-1	1	6.4
1	1	1	1	-1	-1	5.69
1	-1	-1	-1	1	-1	8.82
1	1	-1	-1	1	1	17.55
1	-1	1	-1	1	1	8.87
1	1	1	-1	1	-1	8.94
1	-1	-1	1	1	1	13.06
1	1	-1	1	1	-1	11.49
1	-1	1	1	1	-1	6.25
1	1	1	1	1	1	26.05

2	1	-1	-1	-1	1	5
2	-1	-1	-1	-1	-1	8.11
2	1	1	-1	-1	-1	5.82
2	-1	1	-1	-1	1	7.47
2	1	-1	-1	1	-1	14.23
2	-1	-1	-1	1	1	12.43
2	1	1	-1	1	1	25.38
2	-1	1	-1	1	-1	9.2



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**	
Block	6.81	1	6.81			
Model	764.14	15	50.94	89.46	< 0.0001	significant
A-A	*51.83*	*1*	*51.83*	*91.03*	*< 0.0001*	
B-B	*0.47*	*1*	*0.47*	*0.82*	*0.3942*	
C-C	*2.50*	*1*	*2.50*	*4.38*	*0.0746*	
D-D	*178.72*	*1*	*178.72*	*313.85*	*< 0.0001*	
E-E	*91.03*	*1*	*91.03*	*159.85*	*< 0.0001*	
AB	*42.64*	*1*	*42.64*	*74.88*	*< 0.0001*	
AC	*0.074*	*1*	*0.074*	*0.13*	*0.7298*	
AD	*67.88*	*1*	*67.88*	*119.20*	*< 0.0001*	
AE	*35.19*	*1*	*35.19*	*61.79*	*0.0001*	
BC	*0.073*	*1*	*0.073*	*0.13*	*0.7310*	
BE	*29.02*	*1*	*29.02*	*50.96*	*0.0002*	
DE	*15.80*	*1*	*15.80*	*27.75*	*0.0012*	
ABC	*6.68*	*1*	*6.68*	*11.73*	*0.0111*	
ABE	*34.92*	*1*	*34.92*	*61.32*	*0.0001*	
ADE	*9.58*	*1*	*9.58*	*16.82*	*0.0046*	
Residual	3.99	7	0.57			
Cor Total	774.93	23				

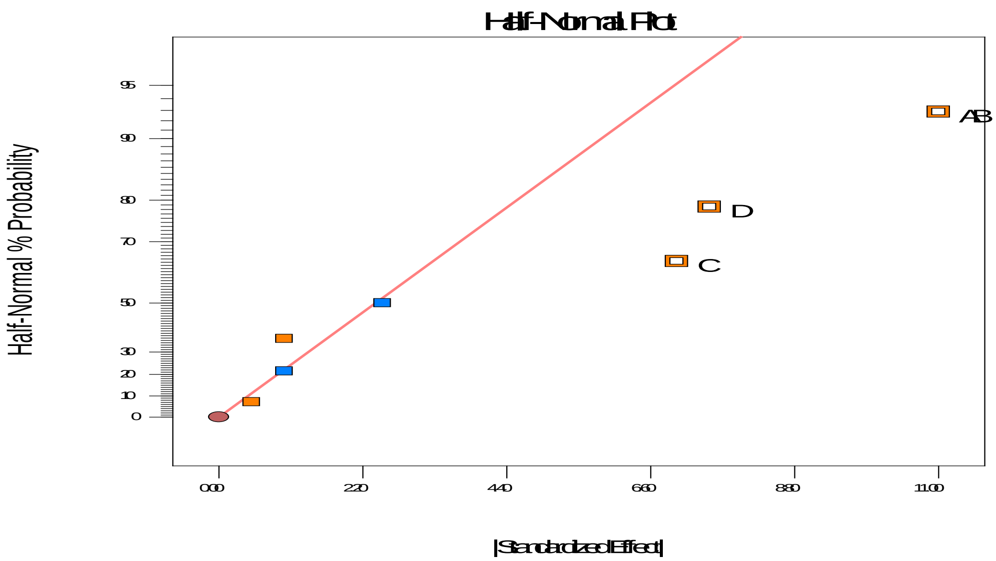
8.33. Consider the 2^4 factorial experiment for surfactin production in Problem 6.40. Suppose that the experimenters could only afford eight runs. Set up the 2^{4-1} fractional factorial with *I* = *ABCD* and select the responses for the runs from the full factorial data in Problem 6.40. Analyze and draw conclusions. Compare your findings with those from the full factorial in Problem 6.40.

The runs for the 2^{4-1} fractional factorial with *I* = *ABCD* are shown below.

Run	Glucose (g dm^{-3})	NH_4NO_3 (g dm^{-3})	$FeSO_4$ (g dm^{-3} x 10^{-4})	$MnSO_4$ (g dm^{-3} x 10^{-2})	y (CMC)$^{-1}$
1	20.00	2.00	6.00	4.00	23
2	60.00	2.00	6.00	20.00	16
3	20.00	6.00	6.00	20.00	18
4	60.00	6.00	6.00	4.00	18
5	20.00	2.00	30.00	20.00	36
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	20.00	34

(a) Analyze the data and draw conclusions.

The analysis below identifies the factors *C* and *D*, and the *AB* interaction as being significant.



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	467.00	5	93.40	74.72	0.0133	significant
A-Glucose	*12.50*	*1*	*12.50*	*10.00*	*0.0871*	
B-NH4NO3	*2.00*	*1*	*2.00*	*1.60*	*0.3333*	
C-FeSO4	*98.00*	*1*	*98.00*	*78.40*	*0.0125*	
D-MnSO4	*112.50*	*1*	*112.50*	*90.00*	*0.0109*	
AB	*242.00*	*1*	*242.00*	*193.60*	*0.0051*	
Residual	2.50	2	1.25			
Cor Total	469.50	7				

(b) Compare your findings with those from the full factorial in Problem 6.40.

In this 2^{4-1} fractional factorial with $I = ABCD$, the AB interaction is aliased with the CD interaction. The full factorial experiment in Problem 6.40 found the CD interaction to be significant. Additional runs, such as a fold-over, should be run to de-alias the CD from the AB interaction.

8.34. Consider the 2^4 factorial experiment in Problem 6.42. Suppose that the experimenters could only afford eight runs. Set up the 2^{4-1}fractional factorial design with $I = ABCD$ and select the responses from the full factorial data in Problem 6.42.

Problem 6.42 was originally analyzed as a replicated 2^3 full factorial experiment. The purpose for this experiment was to compare the fishbone apatite with the hydroxyapatite. However, as shown in the solutions for this problem, the experimenters might have observed better results had they analyzed the experiment as a replicated 2^4 full factorial experiment. This analysis is shown in the solutions for part (i) of Problem 6.42.

The solution presented here is for the replicated 2^{4-1} 2 fractional factorial experiment with $I = ABCD$ as shown below.

Apatite	pH	Pb	Type	Pb,mM	pH
+	+	+	Hydroxyapatite	0.11	3.49
+	+	+	Hydroxyapatite	0.12	3.46
+	+	-	Fishbone	0.01	6.84
+	+	-	Fishbone	0	6.61
+	-	+	Fishbone	1.11	3.35
+	-	+	Fishbone	1.04	3.34
+	-	-	Hydroxyapatite	0.03	3.36
+	-	-	Hydroxyapatite	0.05	3.24
-	+	+	Fishbone	2.11	5.29
-	+	+	Fishbone	2.18	5.06
-	+	-	Hydroxyapatite	0	5.53
-	+	-	Hydroxyapatite	0	5.43
-	-	+	Hydroxyapatite	1.34	2.82
-	-	+	Hydroxyapatite	1.26	2.79
-	-	-	Fishbone	0.05	4.5
-	-	-	Fishbone	0.05	4.74

(a) Analyze the data for all of the responses and draw conclusions.

Due to nonconstant variance observed in the residuals plots, the same transformations that were used in the solution of Problem 6.42 part (i) were also used here.

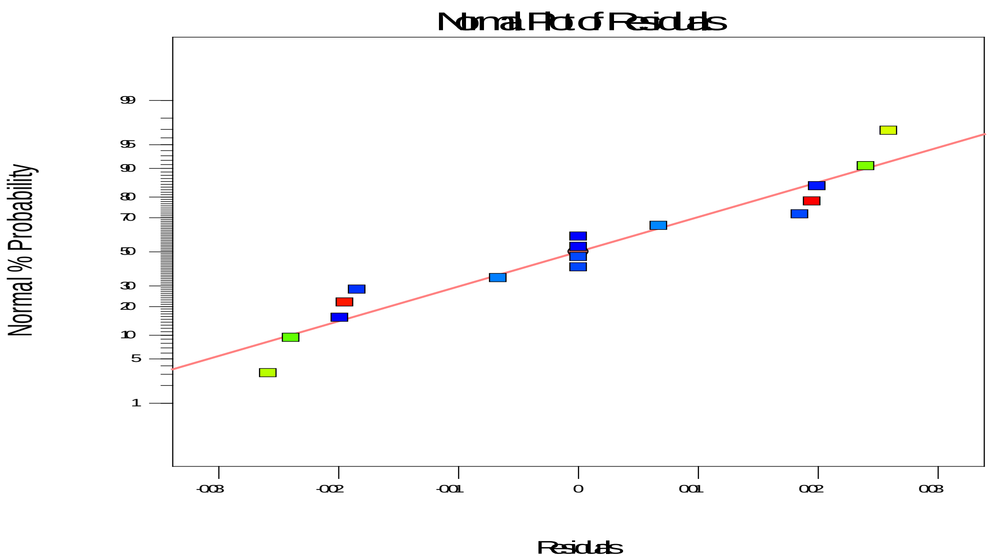
The analysis for the Pb Response identifies all for main effects and the AB, AC, and AD interactions as being significant. However, the AB interaction is aliased with CD, AC is aliased with BD, and AD is aliased with BC.

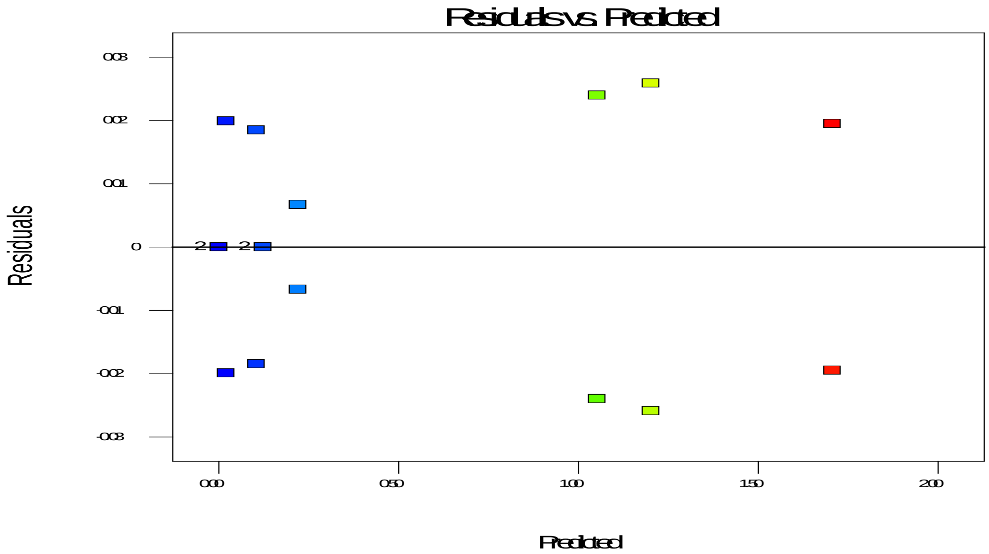
Design Expert Output

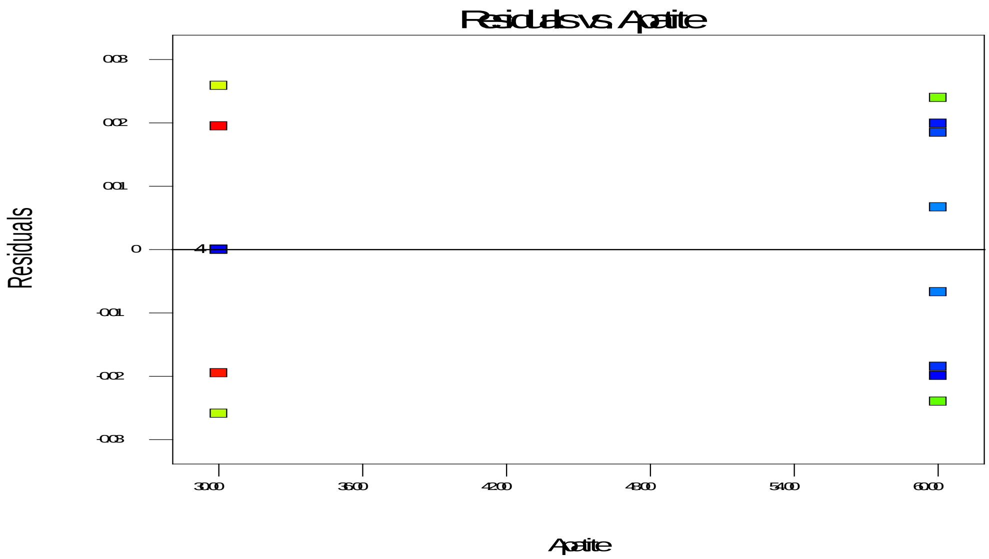
Response **1** **Pb Response**
Transform: **Power** **Lambda:** **0.7** **Constant:** **0**
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

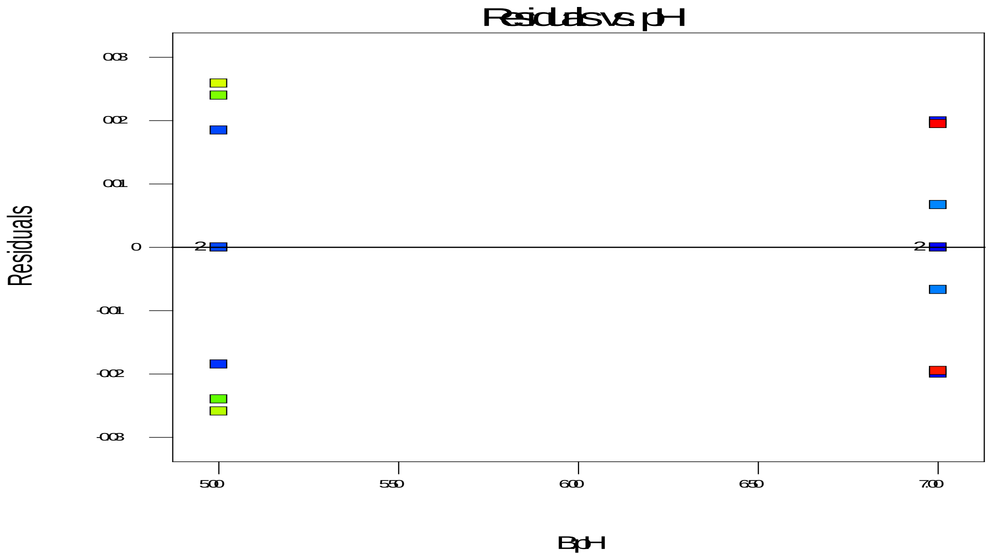
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	6.17	7	0.88	1465.63	< 0.0001	significant
A-Apatite	*0.67*	*1*	*0.67*	*1109.89*	*< 0.0001*	
B-pH	*0.071*	*1*	*0.071*	*118.74*	*< 0.0001*	
C-Pb	*3.87*	*1*	*3.87*	*6425.51*	*< 0.0001*	
D-Type	*0.47*	*1*	*0.47*	*785.36*	*< 0.0001*	
AB	*0.42*	*1*	*0.42*	*700.14*	*< 0.0001*	
AC	*0.67*	*1*	*0.67*	*1113.81*	*< 0.0001*	
AD	*3.600E-003*	*1*	*3.600E-003*	*5.98*	*0.0402*	
Pure Error	4.813E-003	8	6.016E-004			
Cor Total	6.18	15				

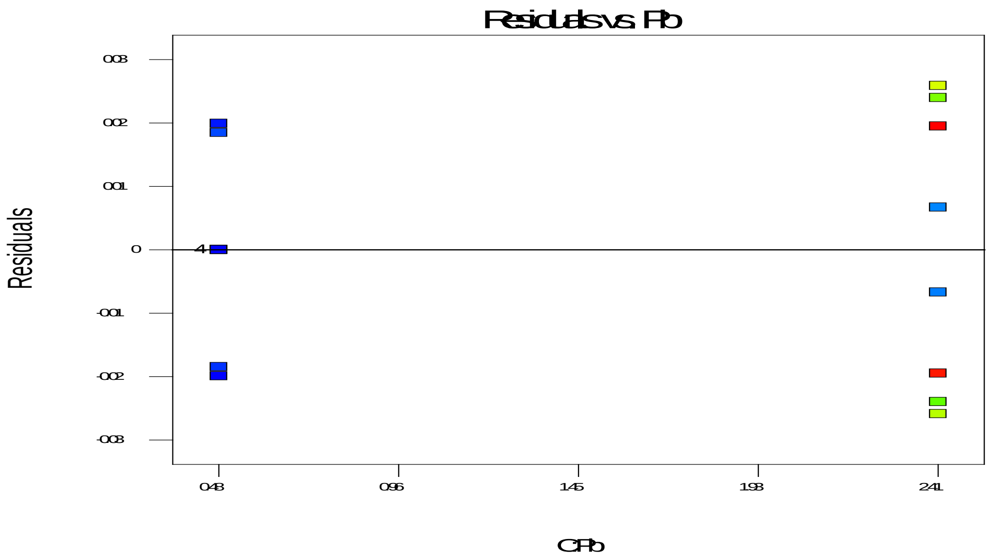
There are no concerns with the residual plots shown below.

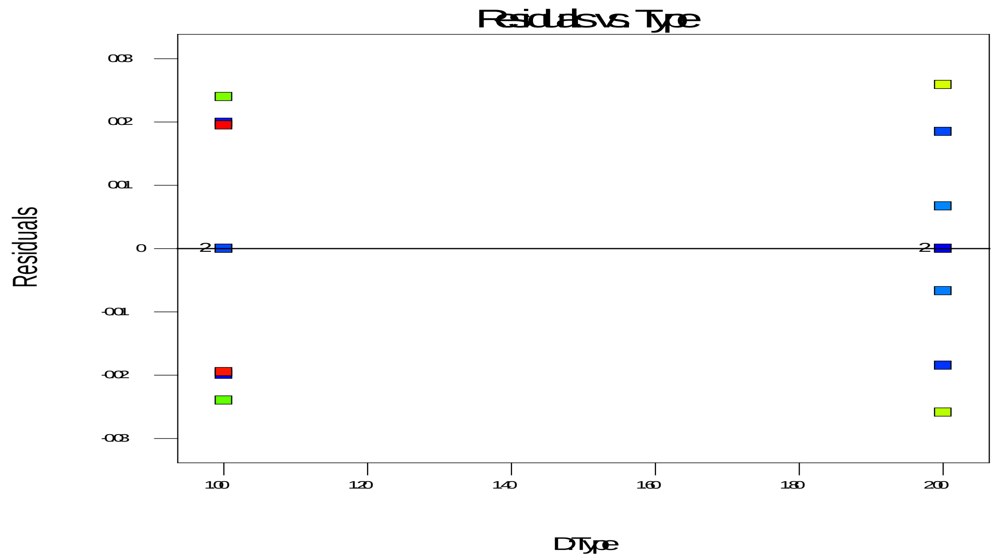



Residuals vs Predicted
Residuals
Predicted


Residuals vs Apatite
Residuals
Apatite


Residuals vs. pH
Residuals
B:pH


Residuals vs. Pb
Residuals
C:Pb



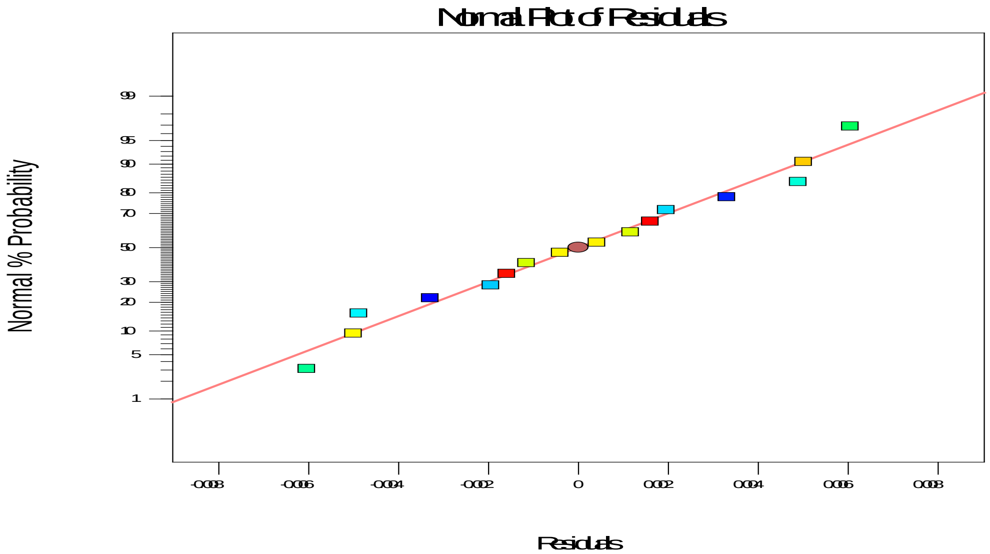
The analysis for the pH Response identifies all for main effects and the *AD* interaction as being significant. However, the AD interaction is aliased with *BC*.

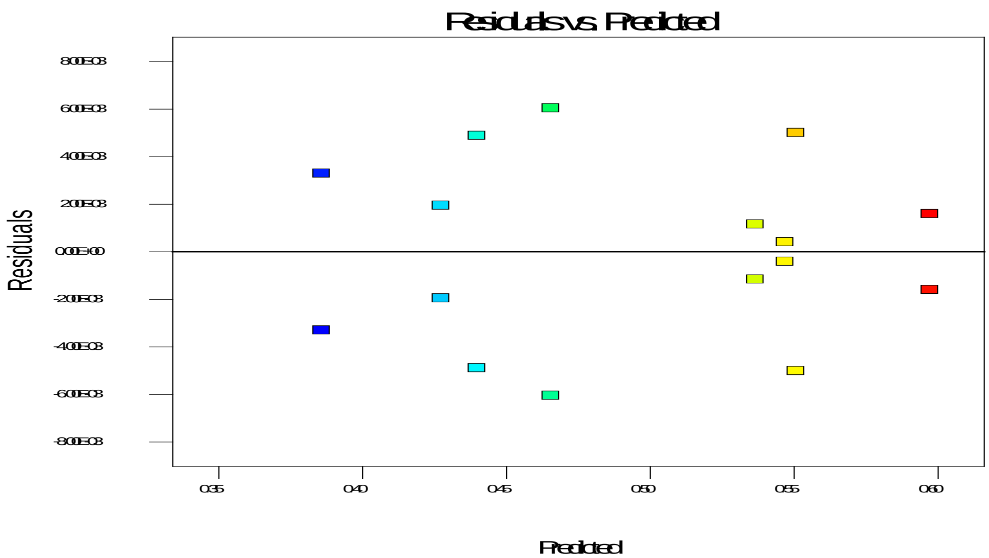
Design Expert Output

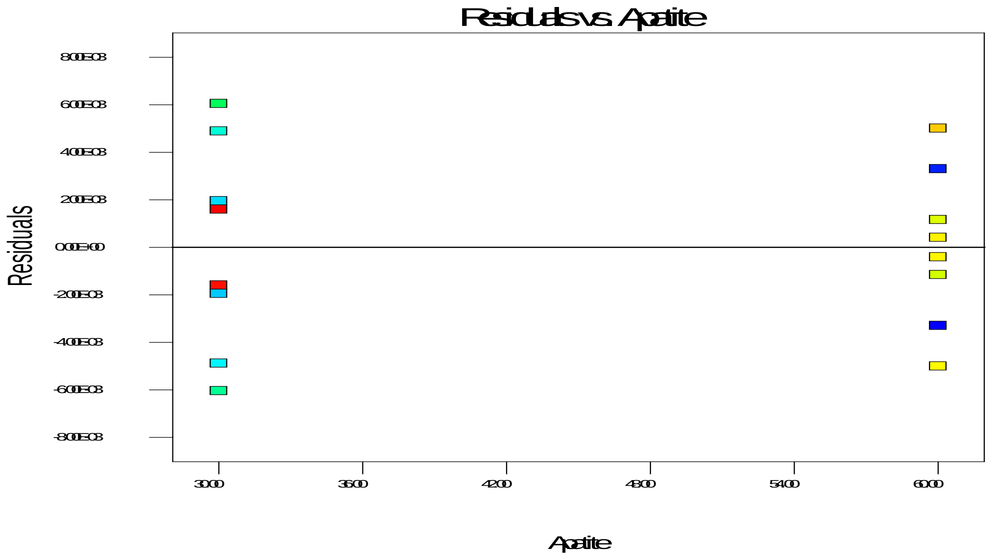
Response **2** **pH Response**
Transform: **Inverse Sqrt** **Constant:** **0**
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

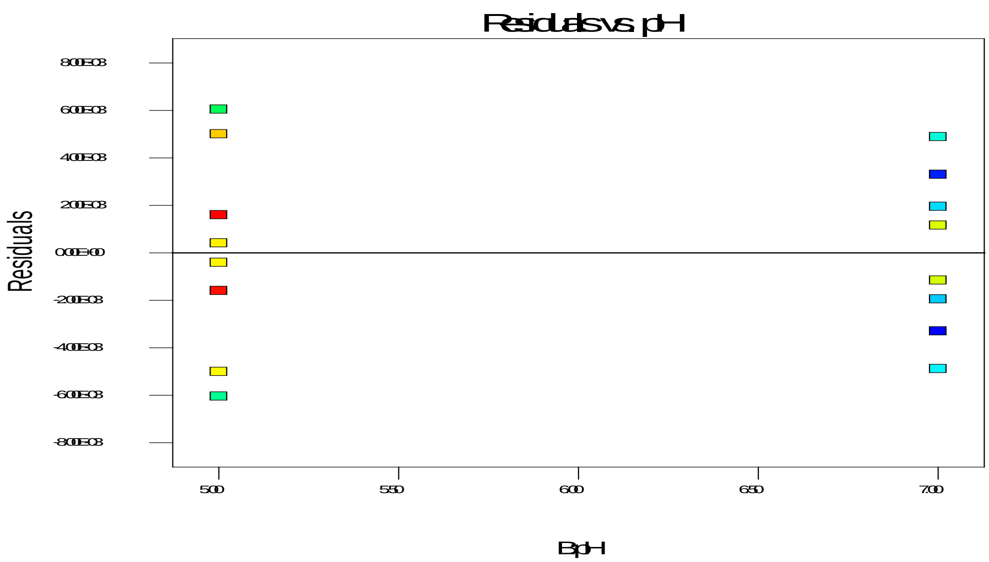
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	0.077	7	0.011	421.02	< 0.0001	significant
A-Apatite	*2.030E-003*	*1*	*2.030E-003*	*77.93*	*< 0.0001*	
B-pH	*0.034*	*1*	*0.034*	*1319.61*	*< 0.0001*	
C-Pb	*0.021*	*1*	*0.021*	*813.91*	*< 0.0001*	
D-Type	*0.019*	*1*	*0.019*	*719.62*	*< 0.0001*	
AB	*1.038E-004*	*1*	*1.038E-004*	*3.98*	*0.0810*	
AC	*1.961E-006*	*1*	*1.961E-006*	*0.075*	*0.7907*	
AD	*3.119E-004*	*1*	*3.119E-004*	*11.97*	*0.0086*	
Pure Error	2.084E-004	8	2.605E-005			
Cor Total	0.077	15				

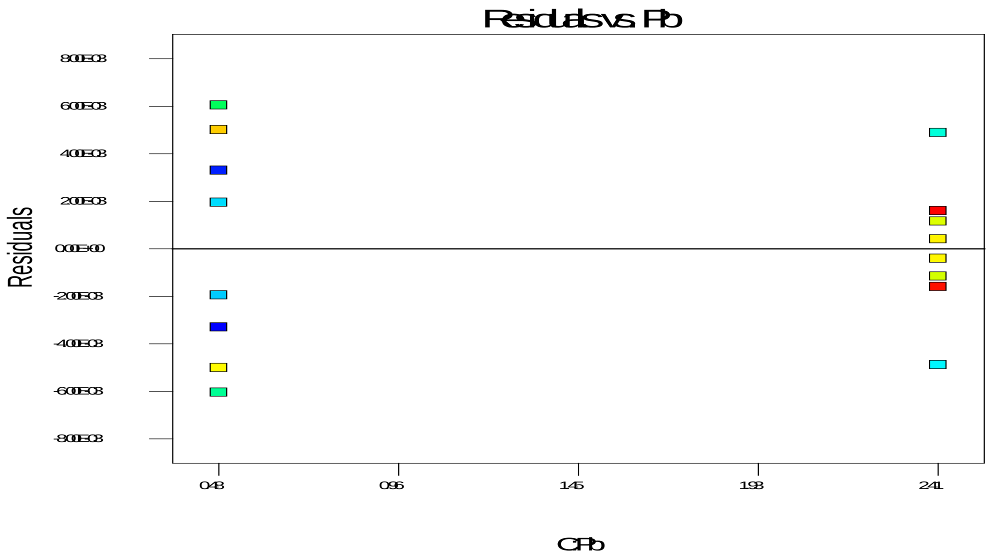
There are no concerns with the residual plots shown below.

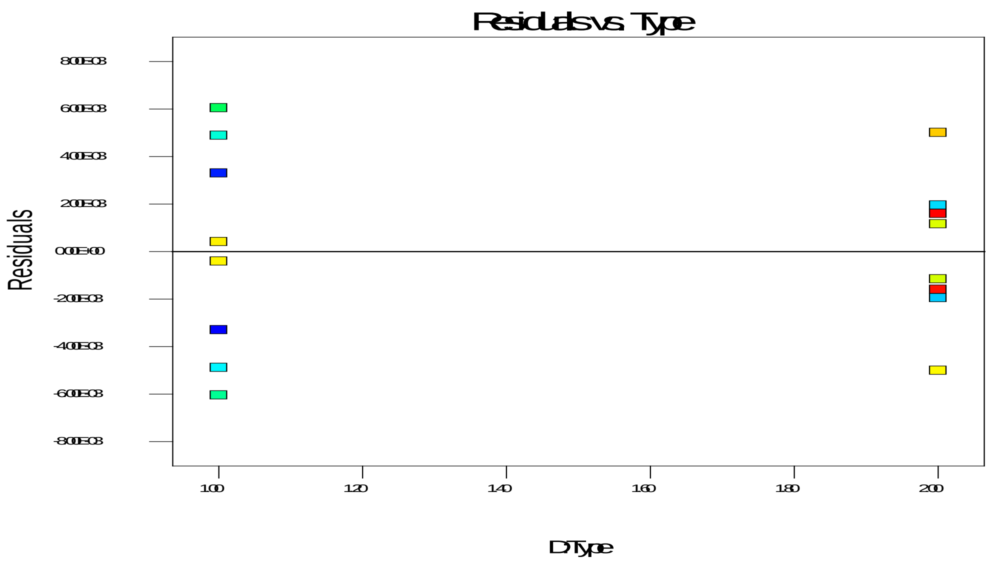

Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
0


Residuals


Residuals vs Apatite
Residuals
Apatite


Residuals vs pH
Residuals





(b) Compare your findings with those from the full factorial in Problem 6.42.

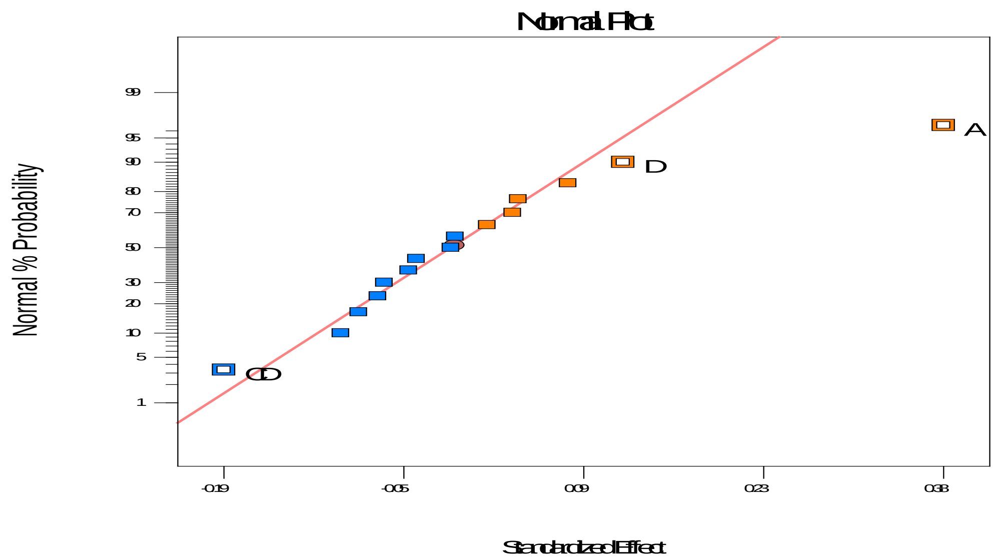
In Problem 6.42, part (i), the main effects, and several two and three factor interactions were found to be very significant. In the 2^{4-1} fractional factorial experiment in Problem 8.34, the main effects are aliased with the three factor interactions, and the two factor interactions are aliased with each other. Based on the results of Problem 6.42, part (i), these aliases are a significant concern.

8.35. An article in the *Journal of Chromatography A* ("Simultaneous Supercritical Fluid Derivatization and Extraction of Formaldehyde by the Hantzsch Reaction," 2000, Vol. 896, pp. 51-59) describes an experiment where the Hantzsch reaction is used to produce the chemical derivatization of formaldehyde in a supercritical medium. Pressure, temperature, and other parameters such as static and dynamic extraction time must be optimized to increase the yield of this kinetically controlled reaction. A 2^{5-1} fractional factorial design with one center run was used to study the significant parameters affecting the supercritical process in terms of resolution and sensitivity. Ultraviolet-visible spectrophotometry was used as the detection technique. The experimental design and the responses are shown in the table below:

Experiment	*P* (MPa)	*T* (°C)	*s* (min)	*d* (min)	*c* (μl)	Resolution	Sensitivity
1	13.8	50	2	2	100	0.00025	0.057
2	55.1	50	2	2	10	0.33333	0.094
3	13.8	120	2	2	10	0.02857	0.017
4	55.1	120	2	2	100	0.20362	1.561
5	13.8	50	15	2	10	0.00027	0.010
6	55.1	50	15	2	100	0.52632	0.673
7	13.8	120	15	2	100	0.00026	0.028
8	55.1	120	15	2	10	0.52632	1.144
9	13.8	50	2	15	10	0.42568	0.142
10	55.1	50	2	15	100	0.60150	0.399
11	13.8	120	2	15	100	0.06098	0.767
12	55.1	120	2	15	10	0.74165	1.086
13	13.8	50	15	15	100	0.08780	0.252
14	55.1	50	15	15	10	0.40000	0.379
15	13.8	120	15	15	10	0.00026	0.028
16	55.1	120	15	15	100	0.28091	3.105
Central	34.5	85	8.5	8.5	55	0.75000	1.836

(a) Analyze the data from this experiment and draw conclusions.

The first analysis shown below is for the Resolution response. Factor *A* and the *CD* interaction are significant with factor *D* being moderately significant. Due to hierarchy, factor *C* is also included in the model. Curvature is also significant.



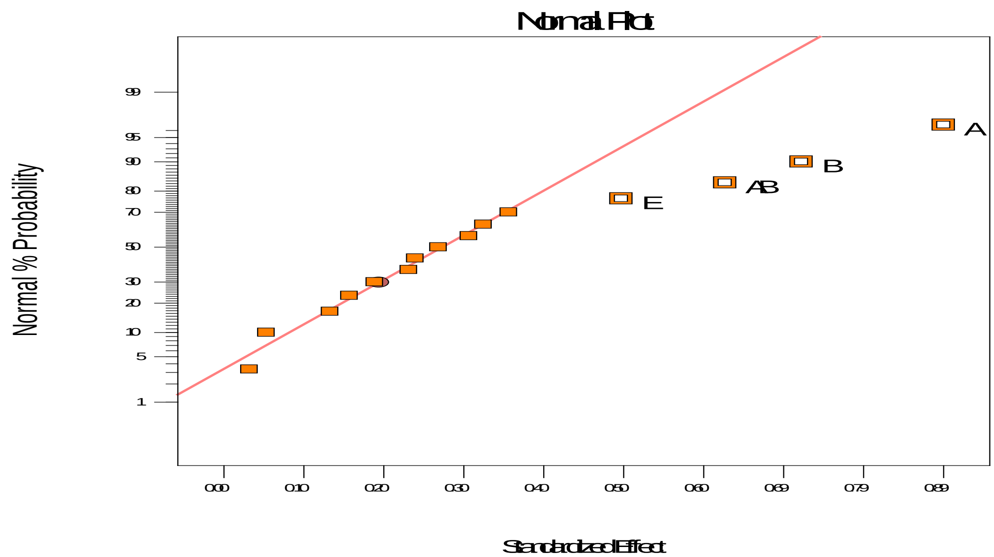
Design Expert Output

Response 1 Resolution
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	0.80	4	0.20	15.37	0.0002	significant
A-P	*0.57*	*1*	*0.57*	*43.70*	*< 0.0001*	
C-s	*0.021*	*1*	*0.021*	*1.59*	*0.2339*	

D-d	*0.060*	*1*	*0.060*	*4.63*	*0.0544*	
CD	*0.15*	*1*	*0.15*	*11.56*	*0.0059*	
Curvature	0.22	1	0.22	17.19	0.0016	significant
Residual	0.14	11	0.013			
Cor Total	1.16	16				

The following analysis is for the Sensitivity response. Factors *A*, *B*, and the *AB* interaction are significant, with factor *E* as being moderately significant. Curvature is also significant.



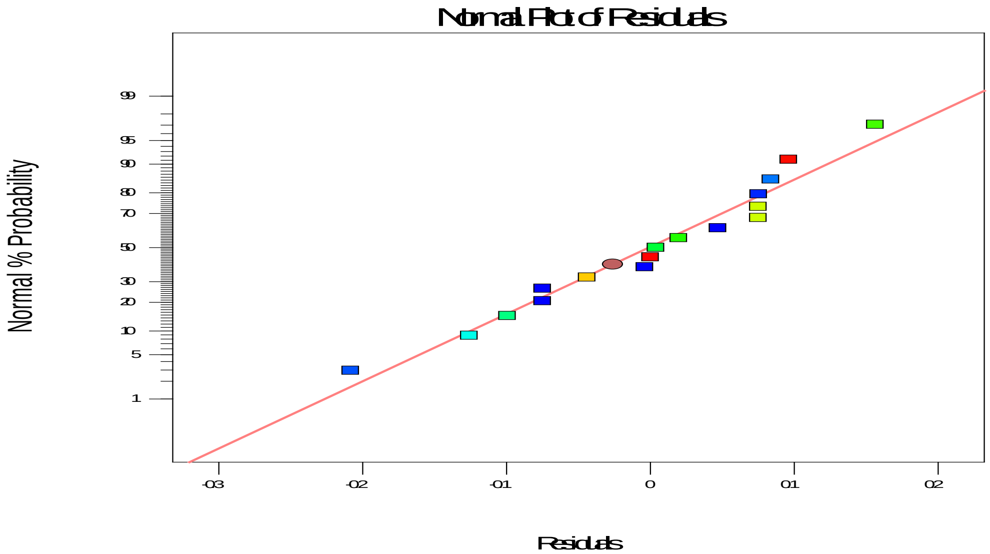
Design Expert Output

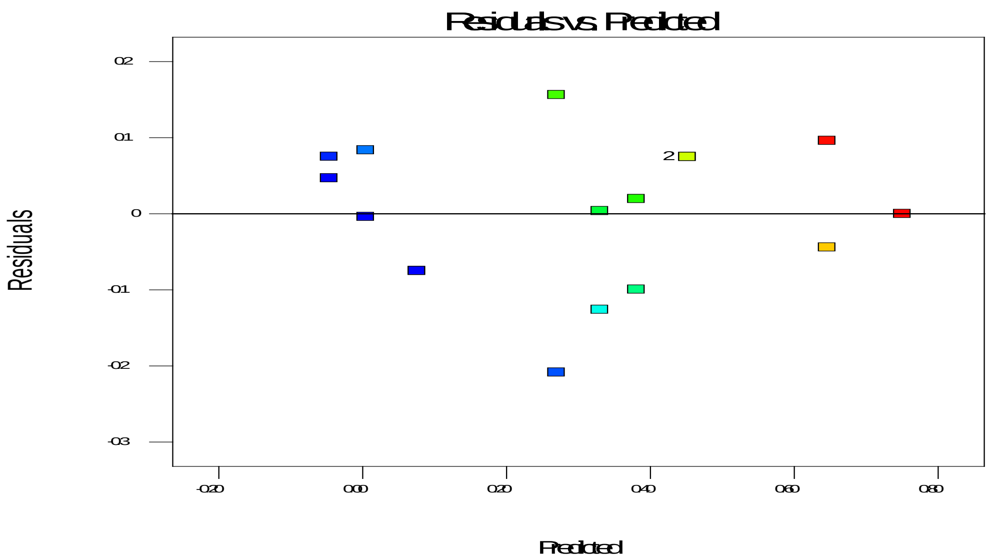
Response 1 Sensitivity
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

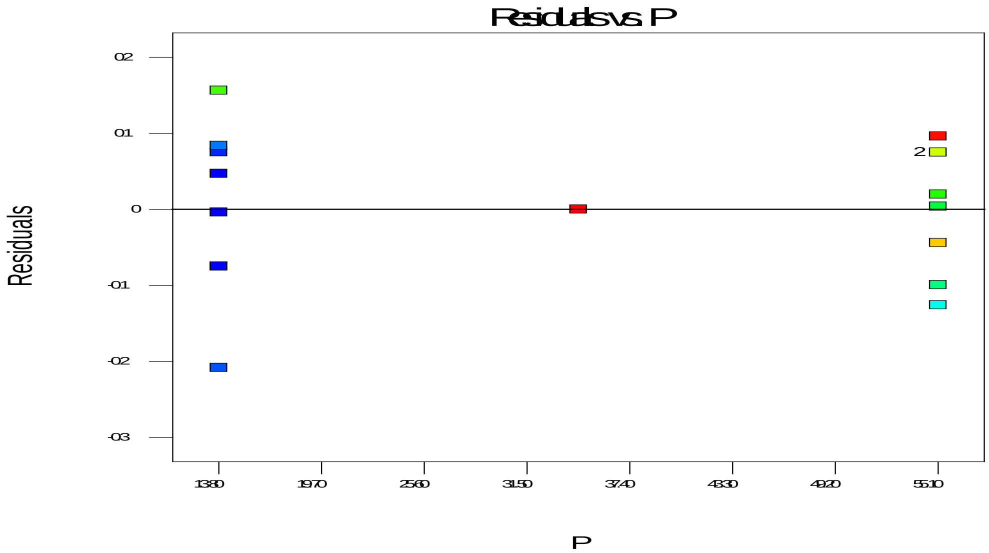
Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Model	7.75	4	1.94	9.19	0.0016	significant
A-P	*3.19*	*1*	*3.19*	*15.10*	*0.0025*	
B-T	*2.05*	*1*	*2.05*	*9.72*	*0.0098*	
E-c	*0.97*	*1*	*0.97*	*4.60*	*0.0551*	
AB	*1.55*	*1*	*1.55*	*7.32*	*0.0204*	
Curvature	1.42	1	1.42	6.72	0.0251	significant
Residual	2.32	11	0.21			
Cor Total	11.49	16				

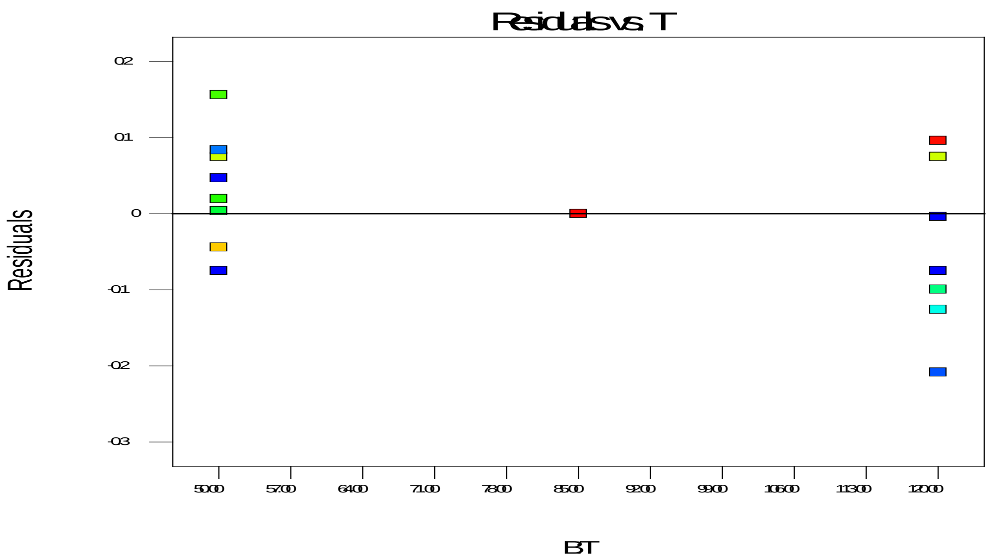
(b) Analyze the residuals. Are there any concerns about model adequacy or violations of the assumptions?

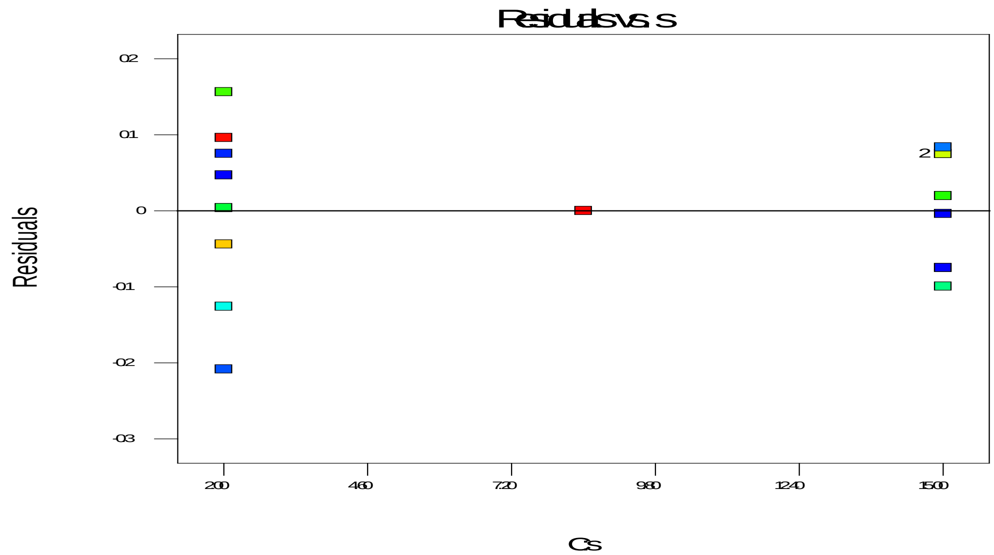
The first set of plots shown below are the residuals for the Resolution response. There does not appear to be any concerns with these residual plots.

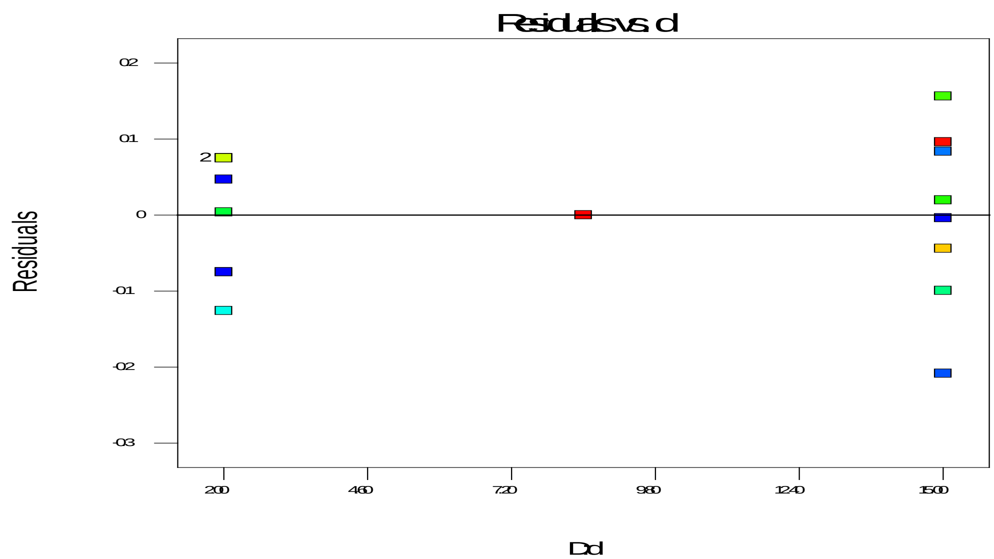

Normal Plot of Residuals
Normal % Probability
Residuals

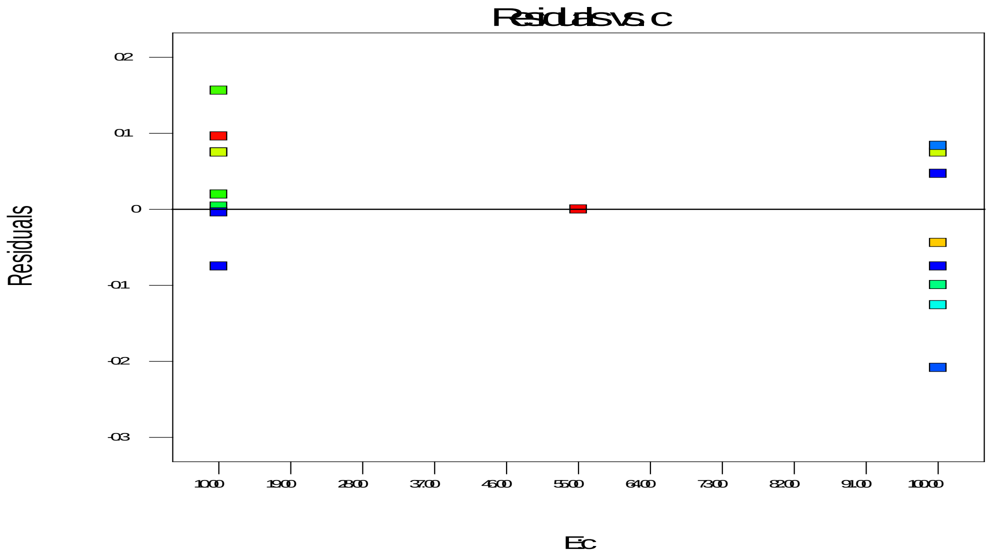

Residuals vs. Predicted
Residuals
Predicted
2


Residuals vs P
Residuals
P
2

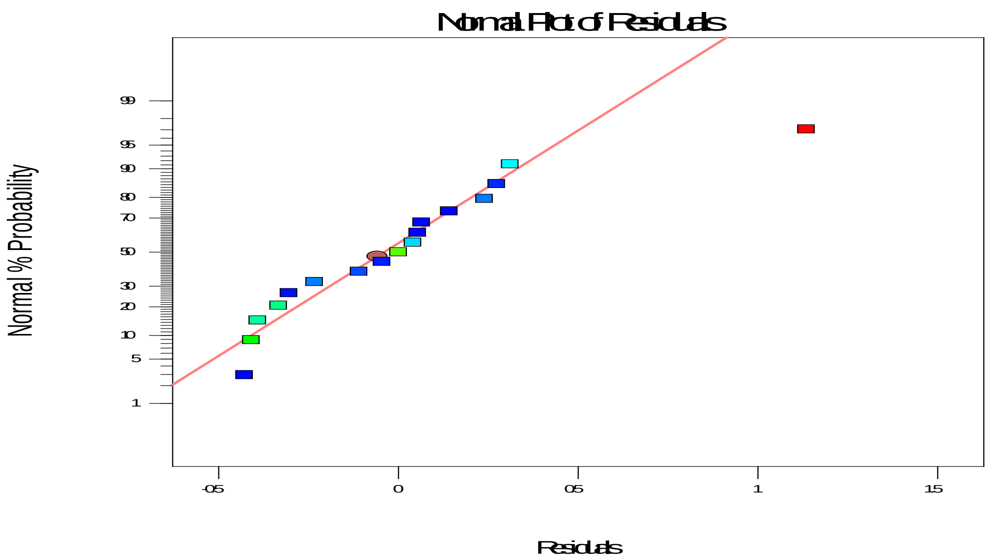

Residuals vs T
Residuals
BT

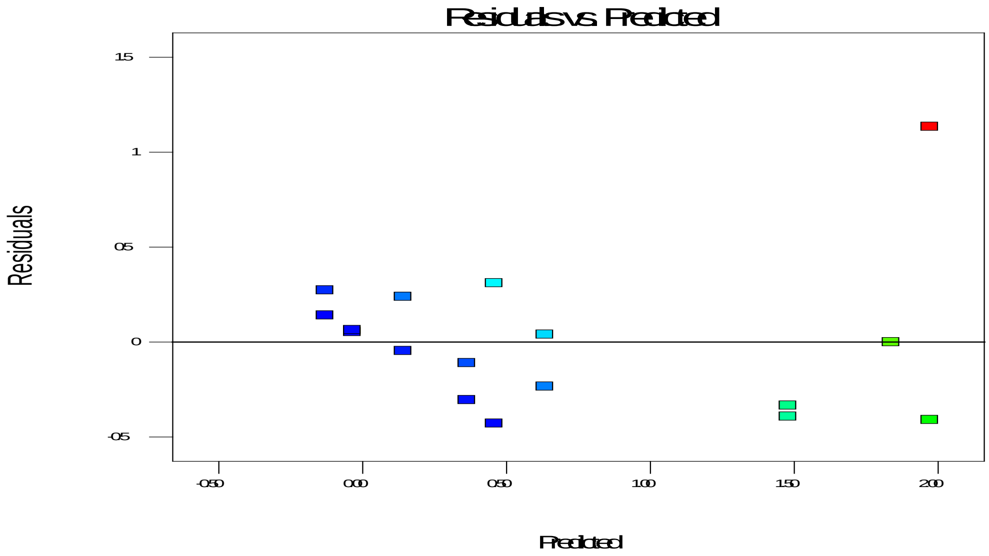

Residuals
2

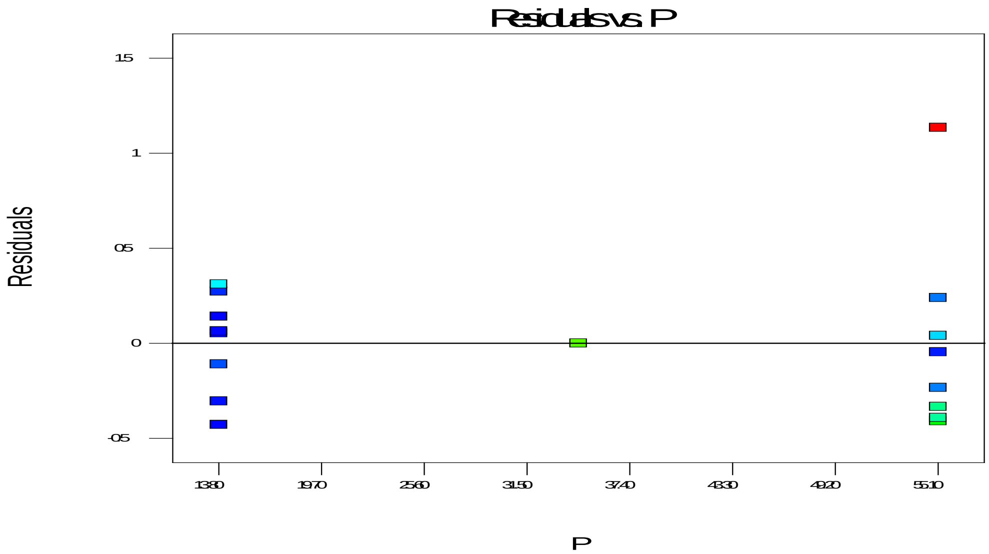

Residuals
2

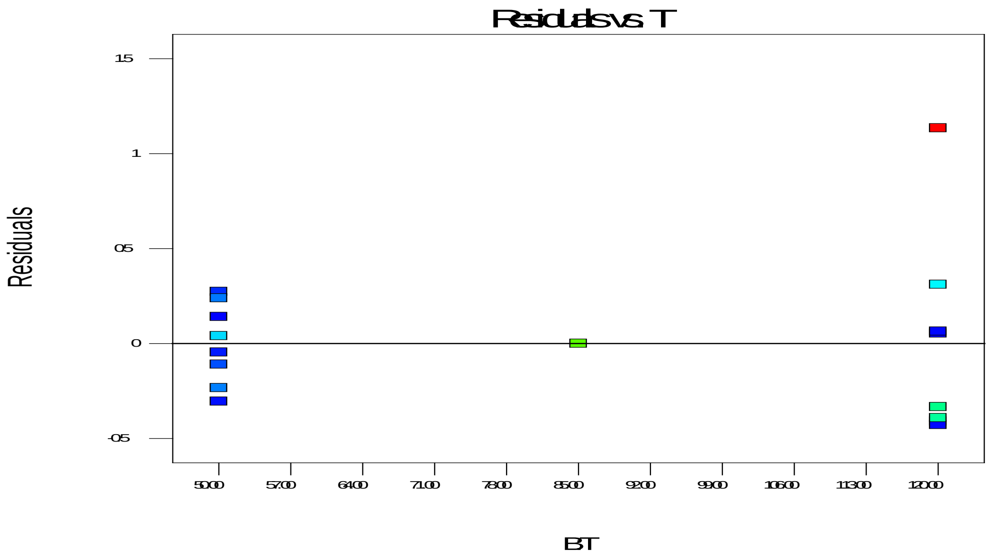


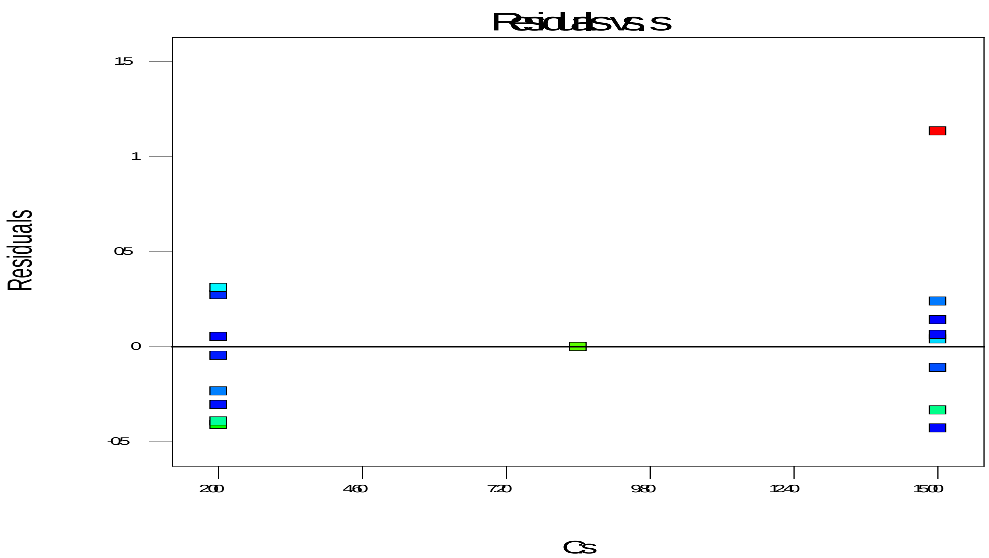
Residual plots for the Sensitivity response are shown below. There appears to be an outlier with the 3.105 Sensitivity value.

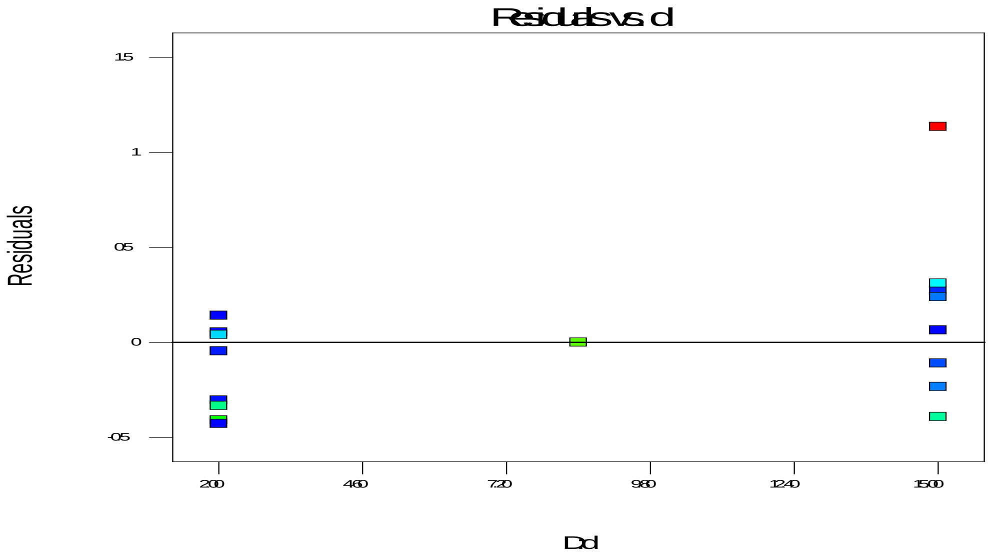


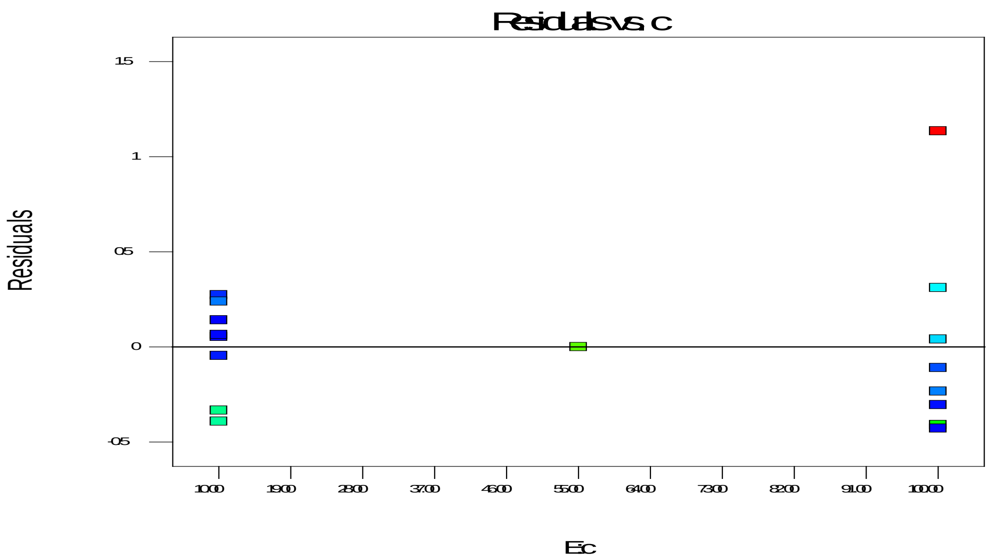

Residuals vs Predicted
Residuals
Predicted


Residuals vs P
Residuals
P


Residuals vs. T
Residuals
B: T


Residuals vs. s
Residuals
C: s





(c) Does the single center point cause any concerns about curvature or the possible need for second-order terms?

Yes, for both Resolution and Sensitivity, there appears to be curvature. Additional runs, such as the axial points for a central composite design, along with additional center points for blocking, could be run to identify the second-order terms.

(d) Do you think that running one center point was a good choice in this design?

It was a good idea to run a center point; however, additional center points would have been an improvement, especially if there was potential need to perform additional runs in a second block to identify the second-order terms.

8.36. An article in *Thin Solid Films* (504, "A Study of Si/SiGe Selective Epitaxial Growth by Experimental Design Approach," 2006, Vol. 504, pp. 95-100) describes the use of a fractional factorial design to investigate the sensitivity of low-temperature (704-760 °C) Si/SiGe selective epitaxial growth to changes in five factors and their two-factor interactions. The five factors are SiH_2Cl_2, GeH_4, HCl, B_2H_6 and temperature. The factor levels studied are:

Factors	Levels (-)	Levels (+)
SiH_2CL_2 (sccm)	8	12
GeH_4 (sccm)	7.2	10.8
HCl (sccm)	3.2	4.8
B_2H_6 (sccm)	4.4	6.6
Temperature (°C)	740	760

The following table contains the design matrix and the three measured responses. Bede RADS Mercury software based on the Takagi-Taupin dynamical scattering theory was used to extract the Si cap thickness, SiGe thickness, and Ge concentration of each sample.

Run Order	Factors A	B	C	D	E	Si cap thickness (Å)	SiGe thickness (Å)	Ge concentration (at. %)
7	-	-	-	-	+	371.18	475.05	8.53
17	-	-	-	+	-	152.36	325.21	9.74
6	-	-	+	-	-	91.69	258.60	9.78
10	-	-	+	+	+	234.48	392.27	9.14
16	-	+	-	-	-	151.36	440.37	12.13
2	-	+	-	+	+	324.49	623.60	10.68
15	-	+	+	-	+	215.91	518.50	11.42
4	-	+	+	+	-	97.91	356.67	12.96
9	+	-	-	-	-	186.07	320.95	7.87
13	+	-	-	+	+	388.69	487.16	7.14
18	+	-	+	-	+	277.39	422.35	6.40
5	+	-	+	+	-	131.25	241.51	8.54
14	+	+	-	-	+	378.41	630.90	9.17
3	+	+	-	+	-	192.65	437.53	10.35
1	+	+	+	-	-	128.99	346.22	10.95
12	+	+	+	+	+	298.40	526.69	9.73
8	0	0	0	0	0	215.70	416.44	9.78
11	0	0	0	0	0	212.21	419.24	9.80

(a) What design did the experimenters use? What is the defining relation?

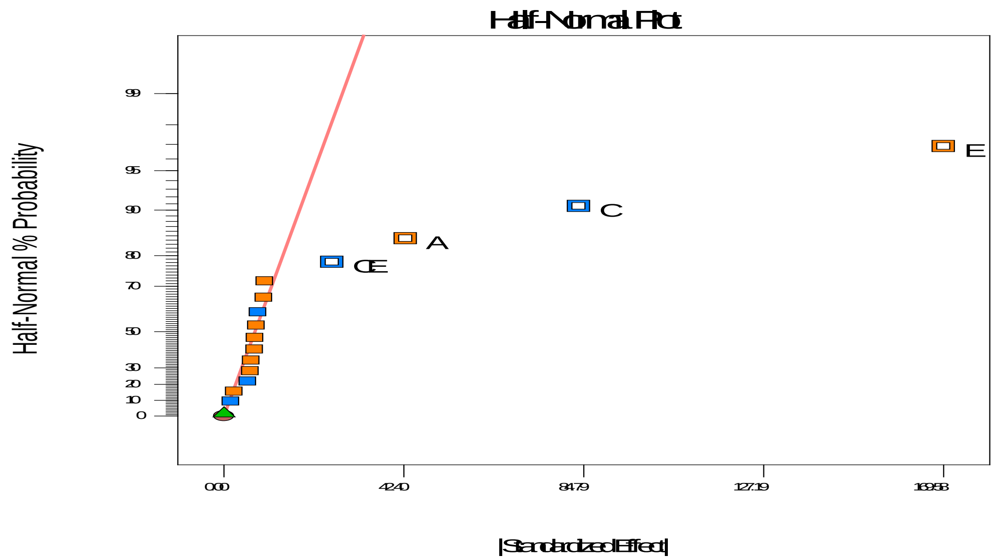
A 2^{5-1} with two center points was run. The defining relation is $I = ABCDE$.

(b) Will the experimenters be able to estimate all main effects and two factor interactions with this experimental design?

Yes, this is a resolution V design. Main factors are confounded with four factor interactions and two factor interactions are confounded with three factor interactions. Assuming sparsity of effects, there is very little chance the three factor and four factor interactions are real.

(c) Analyze all three responses and draw conclusions.

The analysis for the Si cap thickness response is shown below. The *A*, *C*, and *E* factors are significant along with the *CE* interaction.

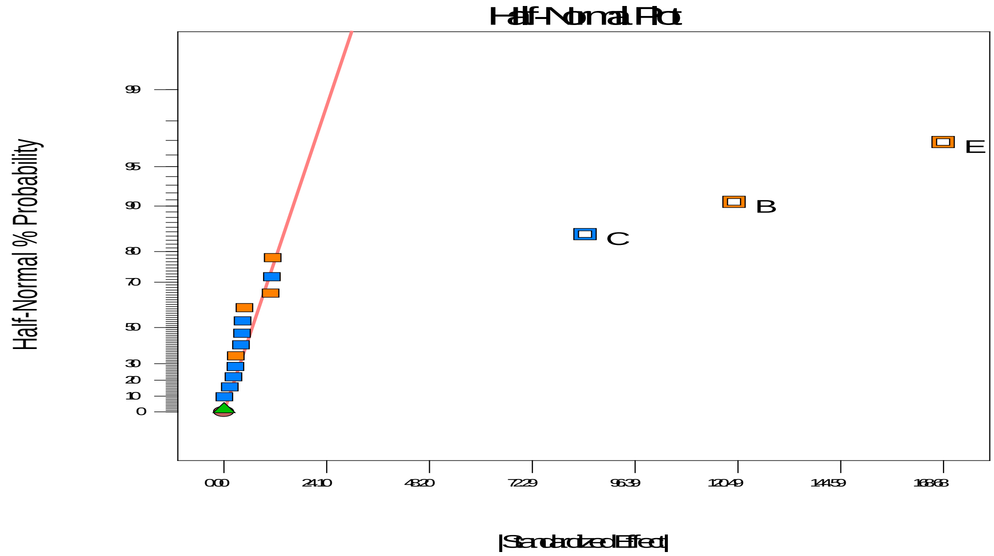


Design Expert Output

Response 1 Si cap thickness
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	1.530E+005	4	38238.54	216.82	< 0.0001	significant
A-SiH2Cl2	*7330.36*	*1*	*7330.36*	*41.56*	*< 0.0001*	
C-HCl	*27988.45*	*1*	*27988.45*	*158.70*	*< 0.0001*	
E-Temperature	*1.150E+005*	*1*	*1.150E+005*	*652.27*	*< 0.0001*	
CE	*2600.75*	*1*	*2600.75*	*14.75*	*0.0024*	
Curvature	272.11	1	272.11	1.54	0.2379	not significant
Residual	2116.33	12	176.36			
Lack of Fit	*2110.24*	*11*	*191.84*	*31.50*	*0.1382*	*not significant*
Pure Error	*6.09*	*1*	*6.09*			
Cor Total	1.553E+005	17				

The analysis for the SiGe thickness response is shown below. The *B*, *C*, and *E* factors are significant. No interactions are significant.

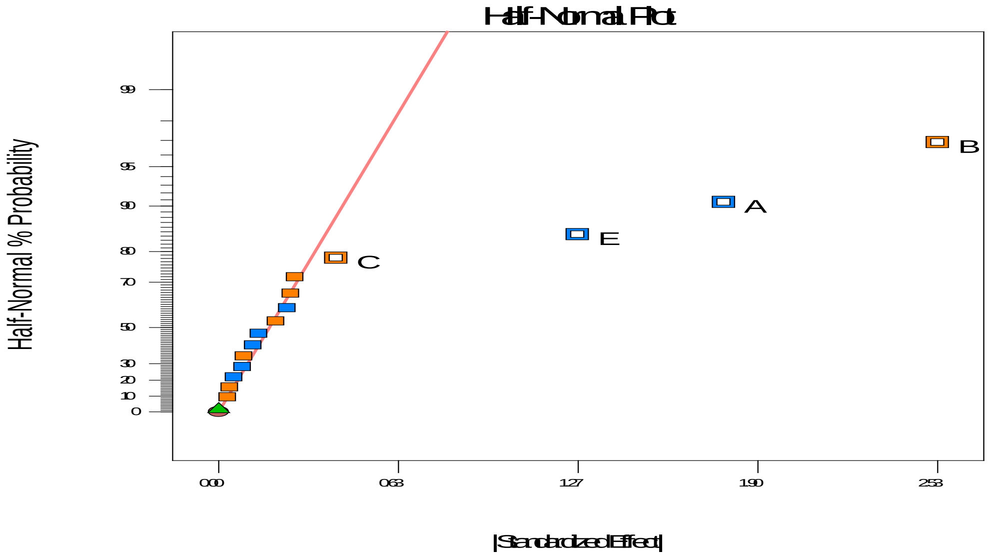


Design Expert Output

Response 1 SiGe thickness
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	1.998E+005	3	66609.34	443.47	< 0.0001	significant
B-GeH4	*57286.03*	*1*	*57286.03*	*381.40*	*< 0.0001*	
C-HCl	*28726.86*	*1*	*28726.86*	*191.26*	*< 0.0001*	
E-Temperature	*1.138E+005*	*1*	*1.138E+005*	*757.75*	*< 0.0001*	
Curvature	96.92	1	96.92	0.65	0.4362	not significant
Residual	1952.61	13	150.20			
Lack of Fit	*1948.69*	*12*	*162.39*	*41.43*	*0.1209*	*not significant*
Pure Error	*3.92*	*1*	*3.92*			
Cor Total	2.019E+005	17				

The analysis for the Ge concentration response is shown below. The *A*, *B*, *C*, and *E* factors are significant. No interactions are significant.



Design Expert Output

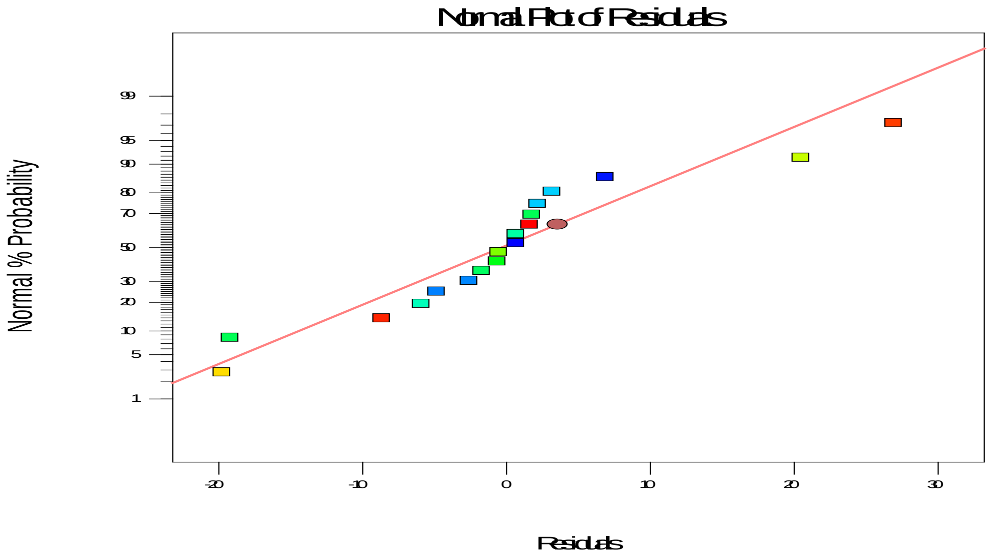
Response 1 Ge concentration
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

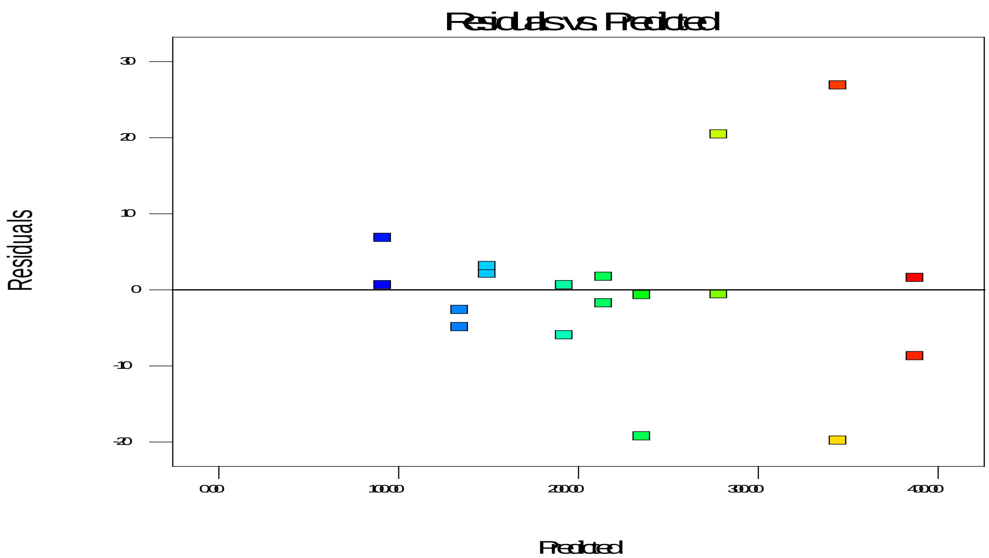
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	45.36	4	11.34	117.19	< 0.0001	significant
A-SiH2Cl2	*12.66*	*1*	*12.66*	*130.80*	*< 0.0001*	
B-GeH4	*25.63*	*1*	*25.63*	*264.87*	*< 0.0001*	
C-HCl	*0.68*	*1*	*0.68*	*7.08*	*0.0208*	
E-Temperature	*6.39*	*1*	*6.39*	*66.02*	*< 0.0001*	
Curvature	0.031	1	0.031	0.32	0.5823	not significant
Residual	1.16	12	0.097			
Lack of Fit	*1.16*	*11*	*0.11*	*527.69*	*0.0339*	*significant*
Pure Error	*2.000E-004*	*1*	*2.000E-004*			
Cor Total	46.55	17				

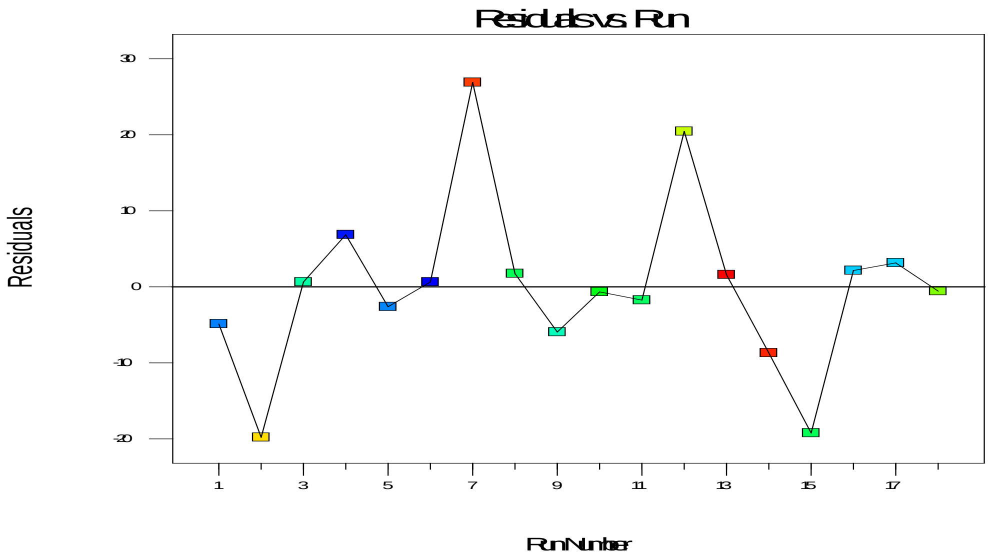
(d) Is there any indication of curvature in the responses?

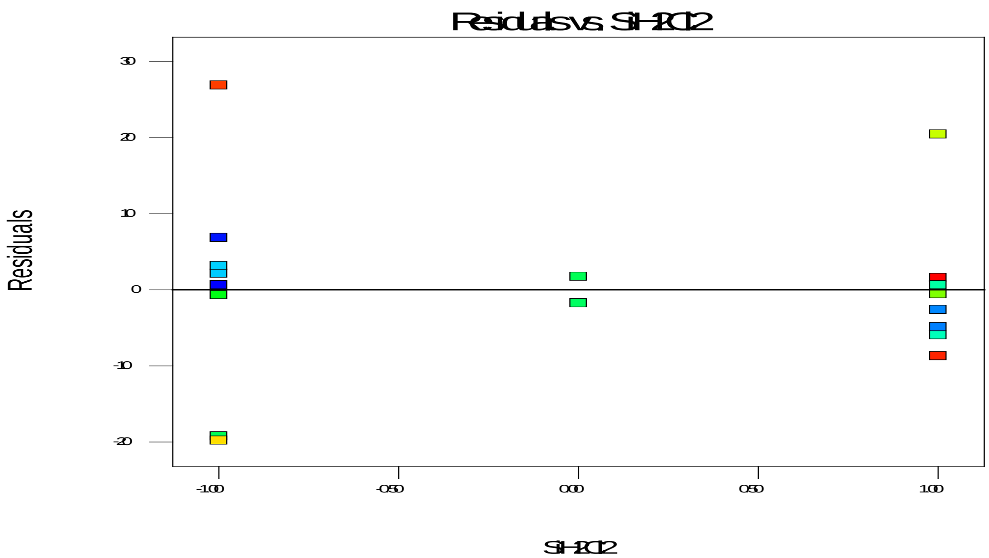
(e) Analyze the residuals and comment on model adequacy.

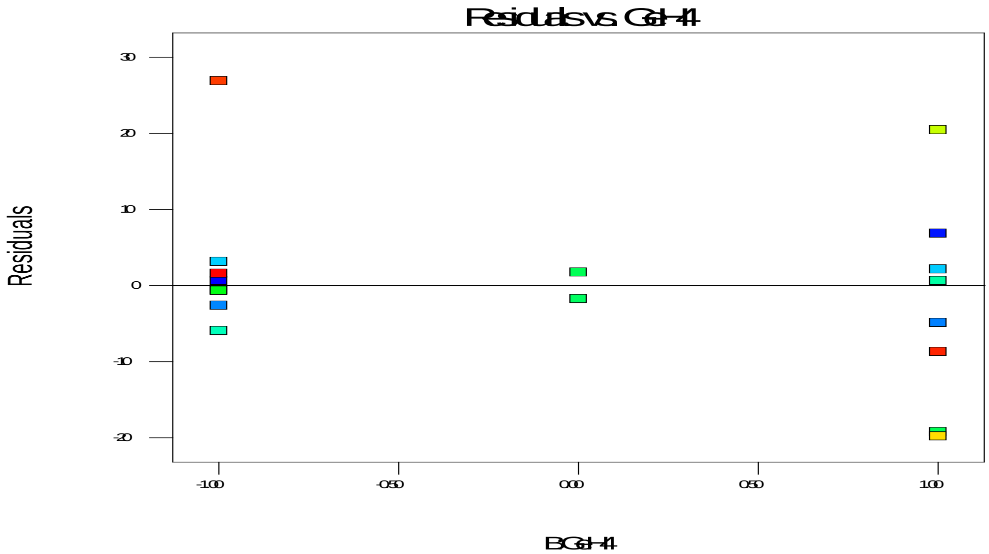
The residual plots for the Si cap thickness response are shown below. There is a concern with inequality of variance. A natural log transformation was applied with the analysis and residual plots shown at the end of this problem.

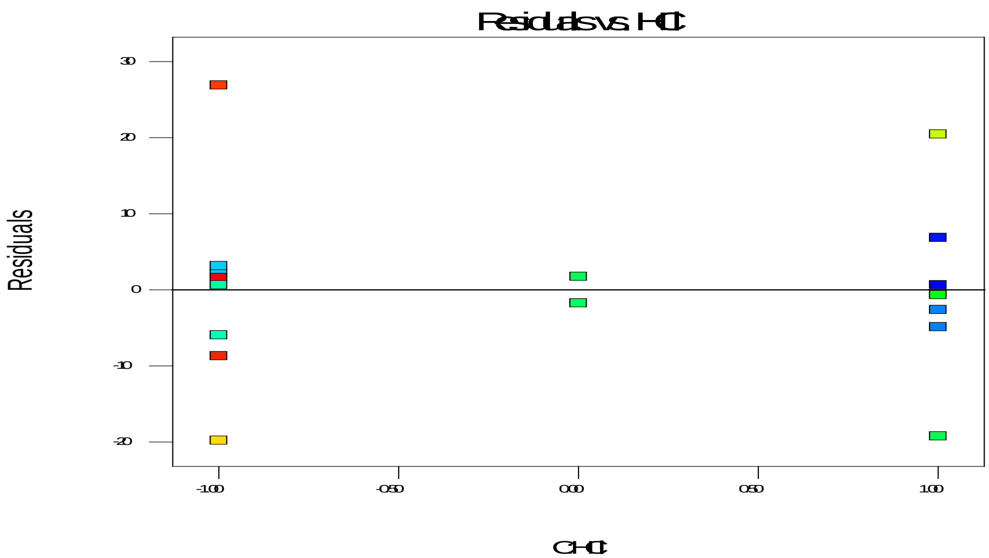

Normal Plot of Residuals
Normal % Probability
Residuals

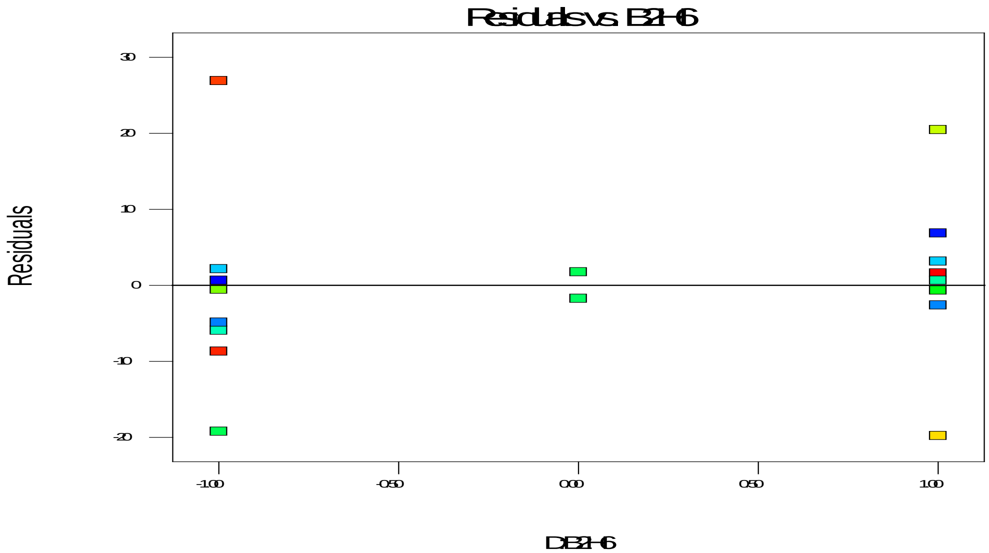

Residuals vs. Predicted
Residuals
Predicted

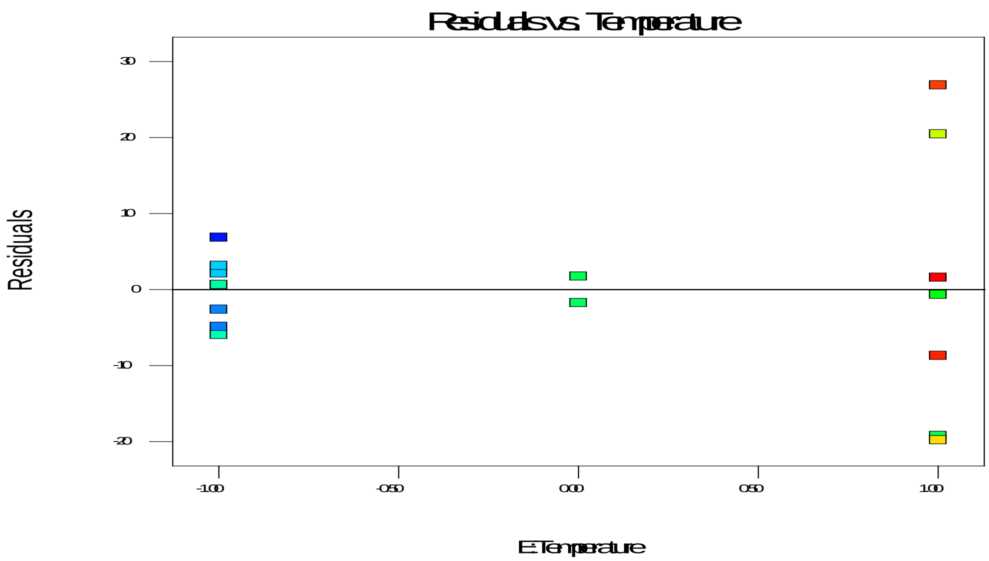

Residuals vs Run
30
20
10
0
-10
-20
Residuals
1
3
5
7
9
11
13
15
17


30
20
10
0
-10
-20
Residuals

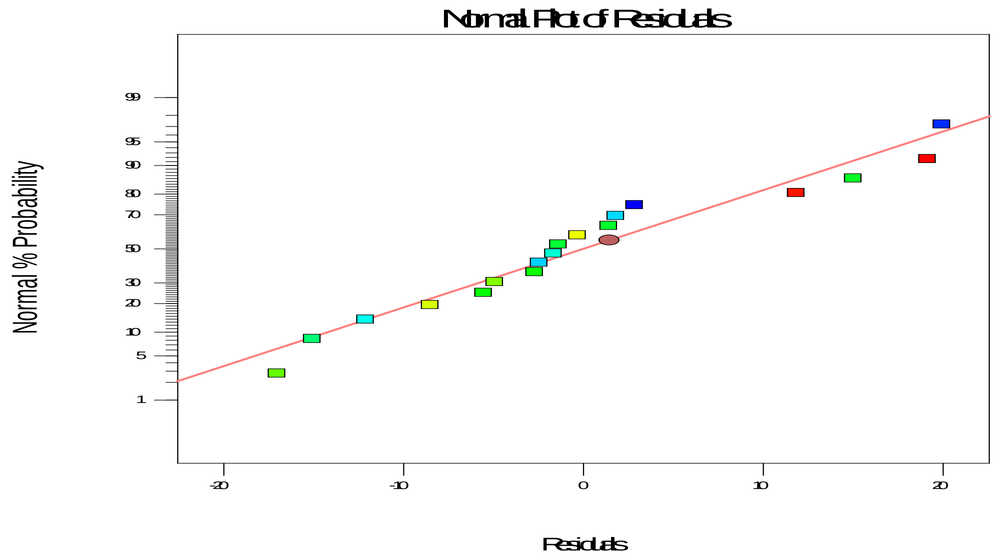

Residuals vs. C6H14
Residuals
30
20
10
0
-10
-20
-1.00
-0.50
0.00
0.50
1.00
B:C6H14

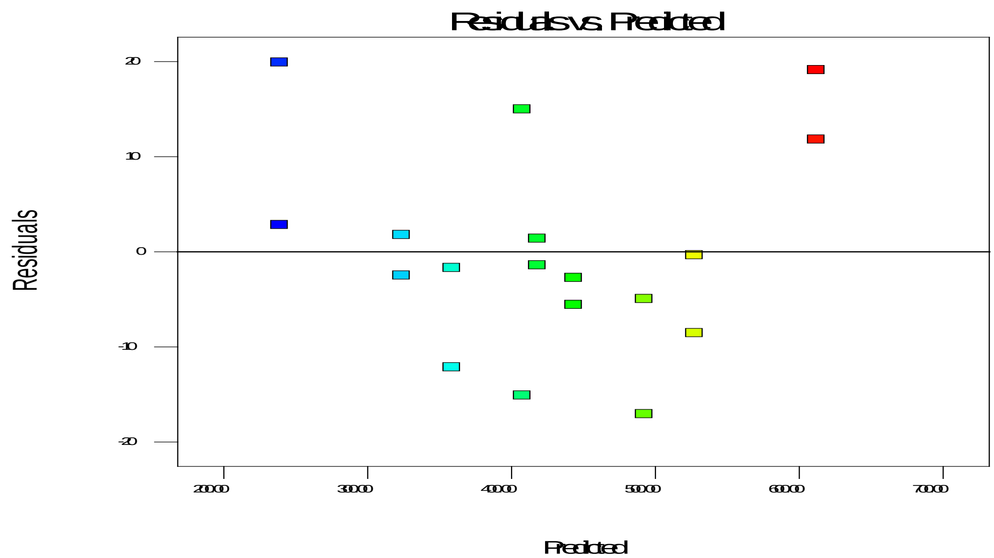

Residuals vs. HCl
Residuals
30
20
10
0
-10
-20
-1.00
-0.50
0.00
0.50
1.00
C:HCl

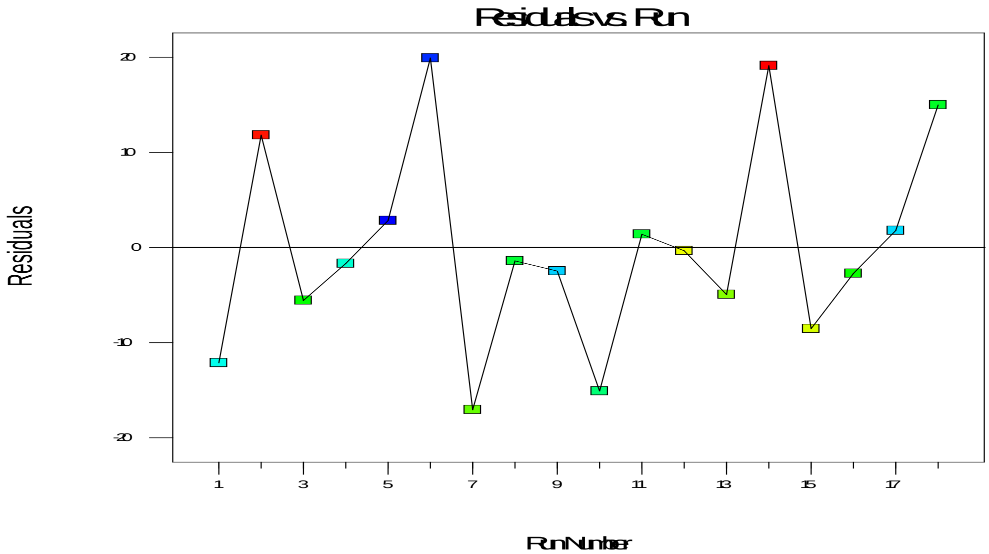


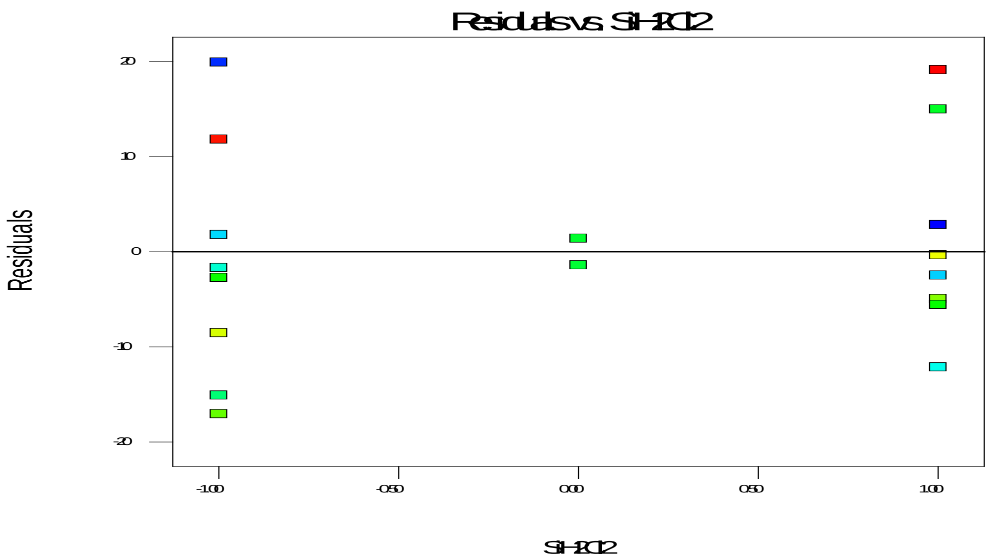


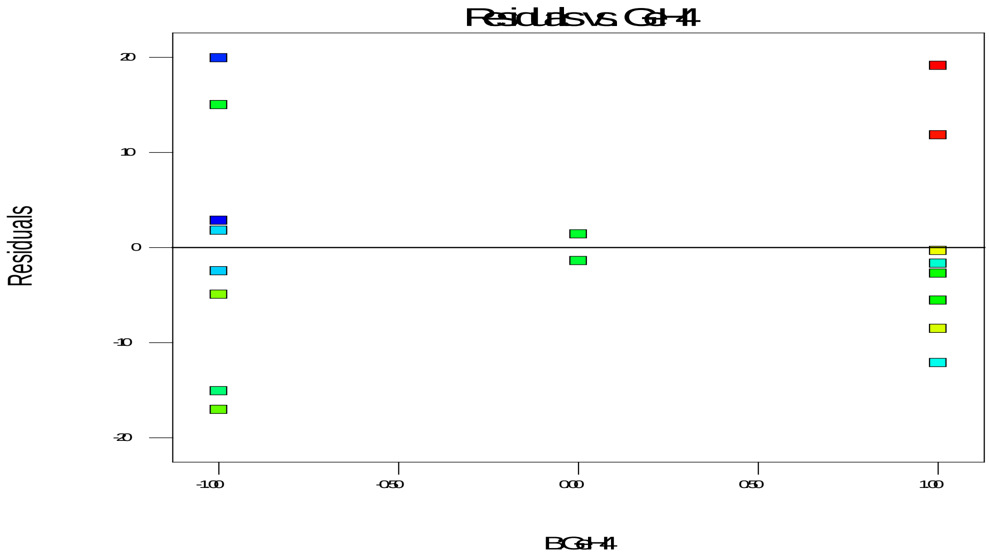
The residual plots for the SiGe thickness response are shown below. There are no concerns with these residual plots.

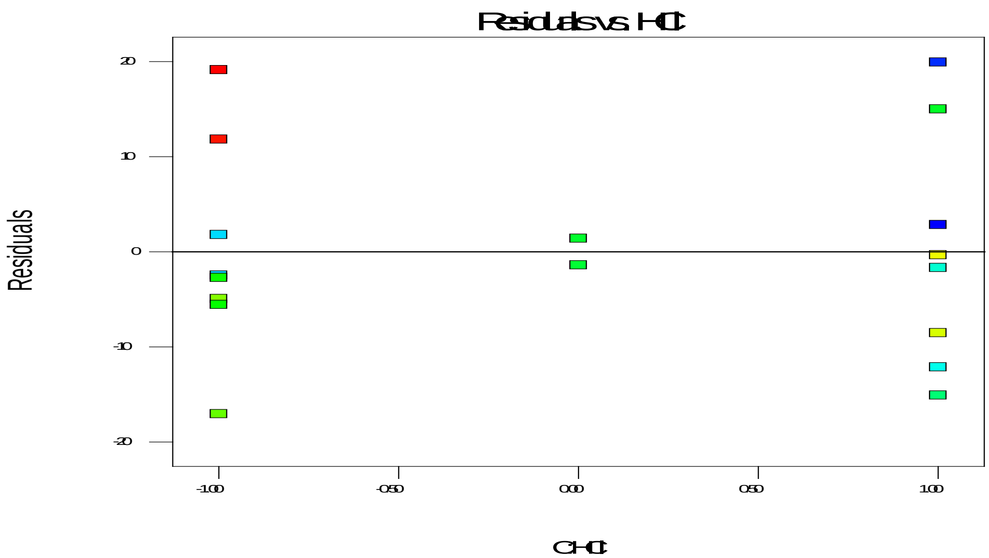

Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-20
-10
0
10
20
Residuals

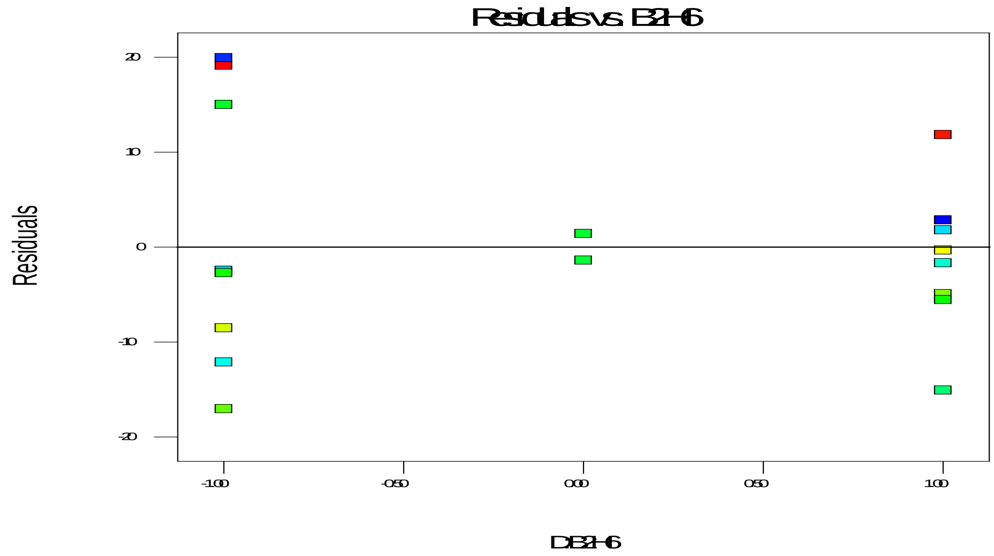

Residuals vs. Predicted
Residuals
20
10
0
-10
-20
2000
3000
4000
5000
6000
7000
Predicted

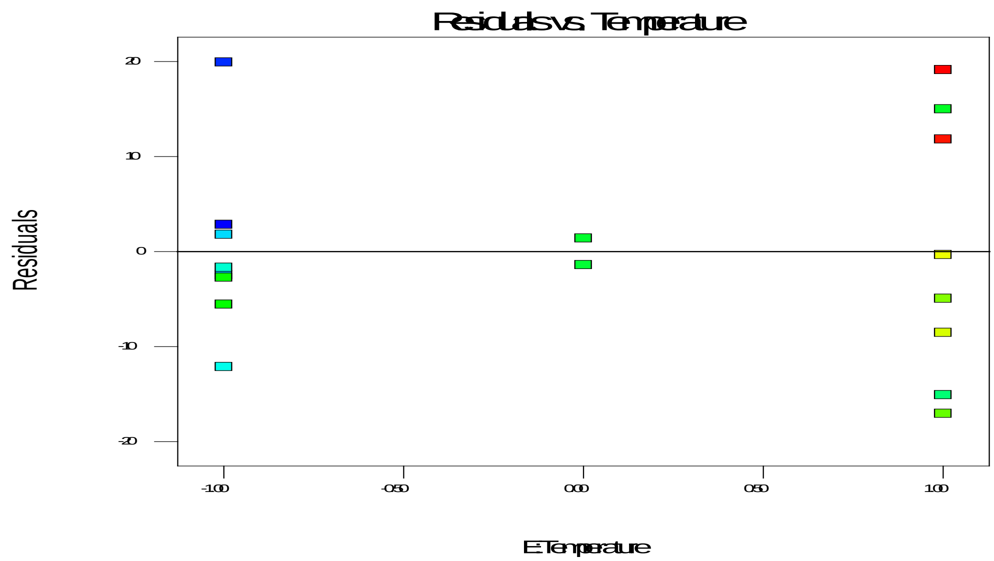

Residuals vs Run
Residuals
Run Number


Residuals vs SiH2Cl2
Residuals
SiH2Cl2

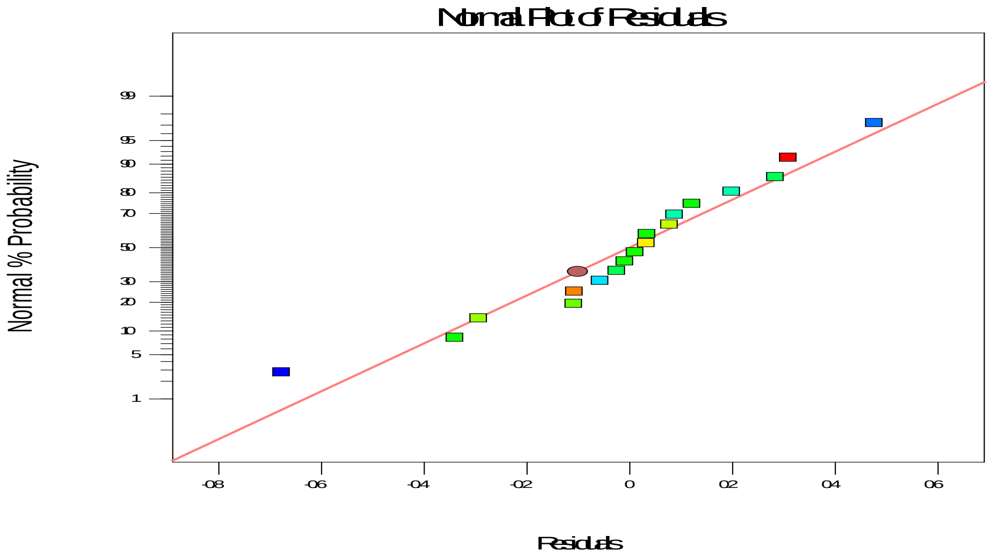

Residuals
20
10
0
-10
-20
-1.00
-0.50
0.00
0.50
1.00

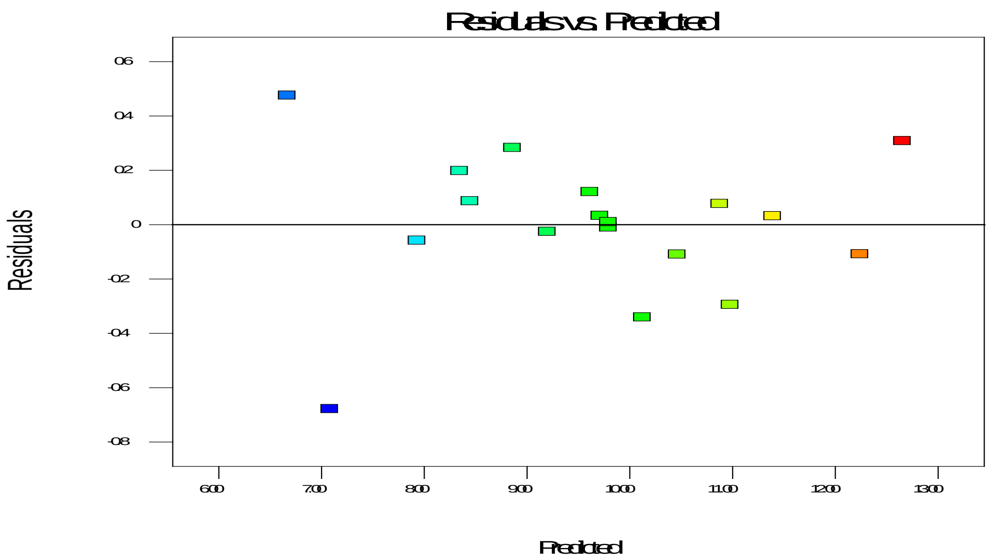

Residuals
20
10
0
-10
-20
-1.00
-0.50
0.00
0.50
1.00

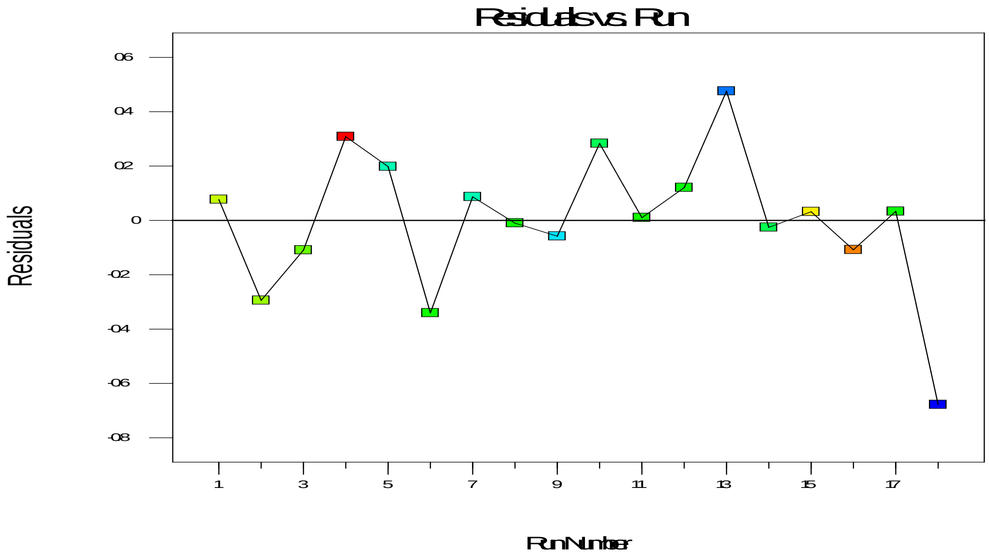


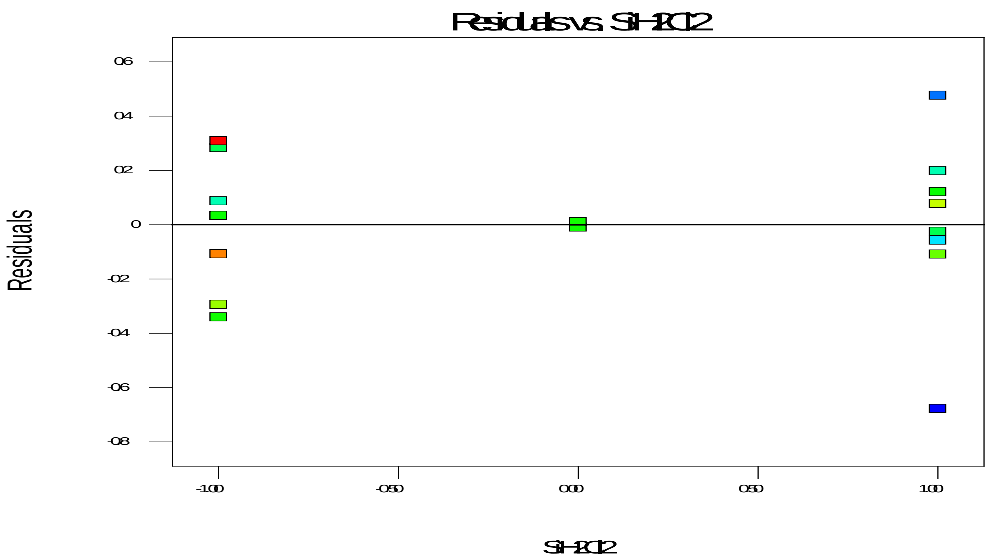


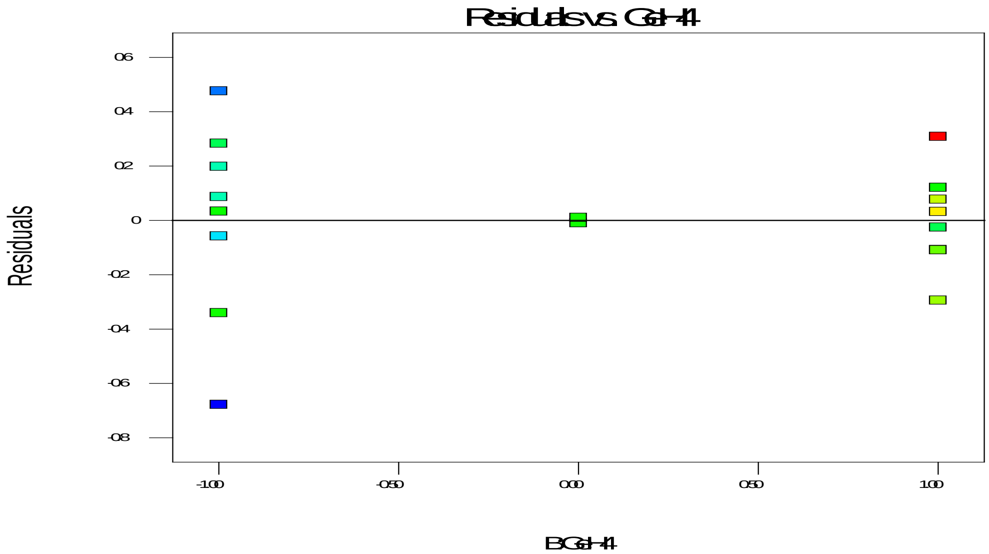
The residual plots for the Ge concentration response are as follows. Run number 18, Ge concentration of 6.40, appears to be a slight outlier. Additional investigation by the experimenter may be necessary.

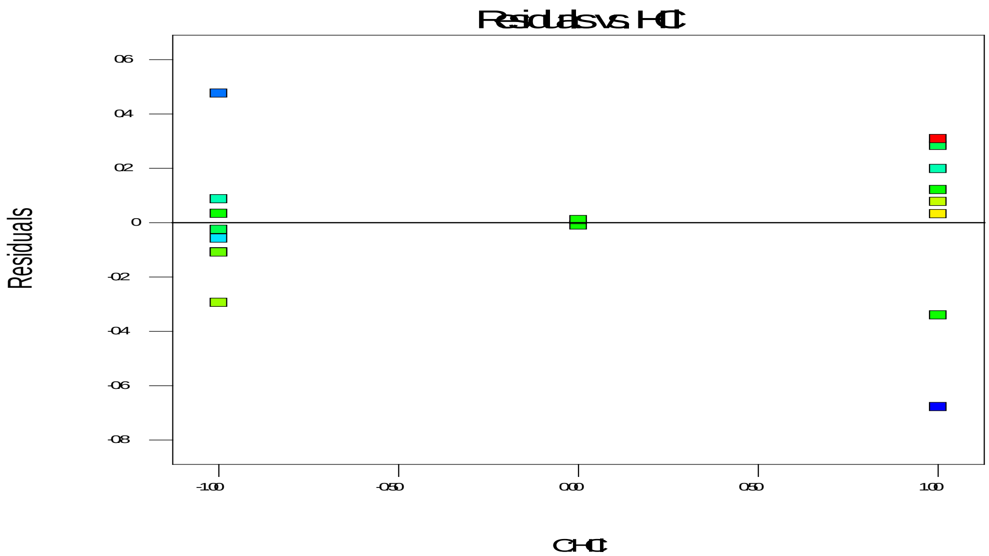

Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
Residuals

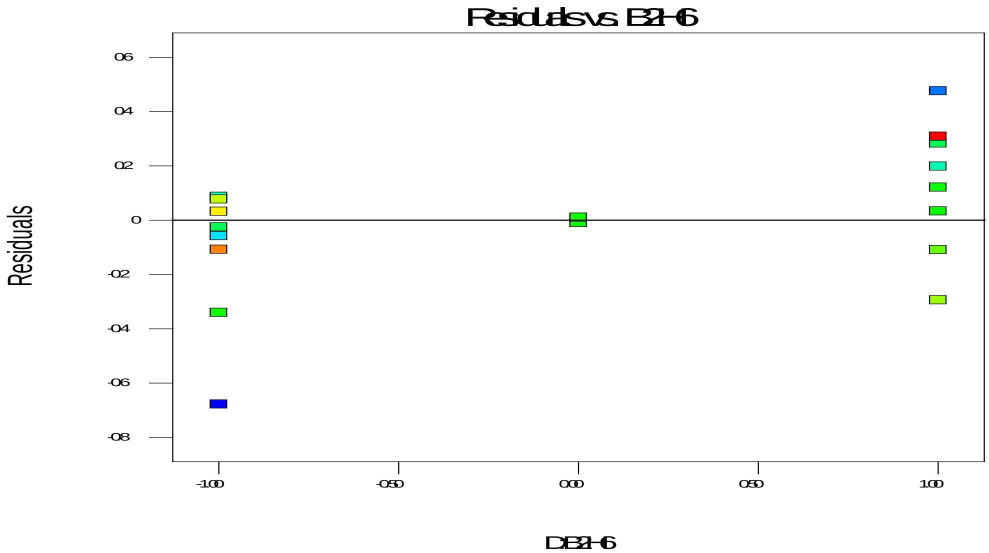

Residuals vs. Predicted
Residuals
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
60.0
70.0
80.0
90.0
100.0
110.0
120.0
130.0
Predicted

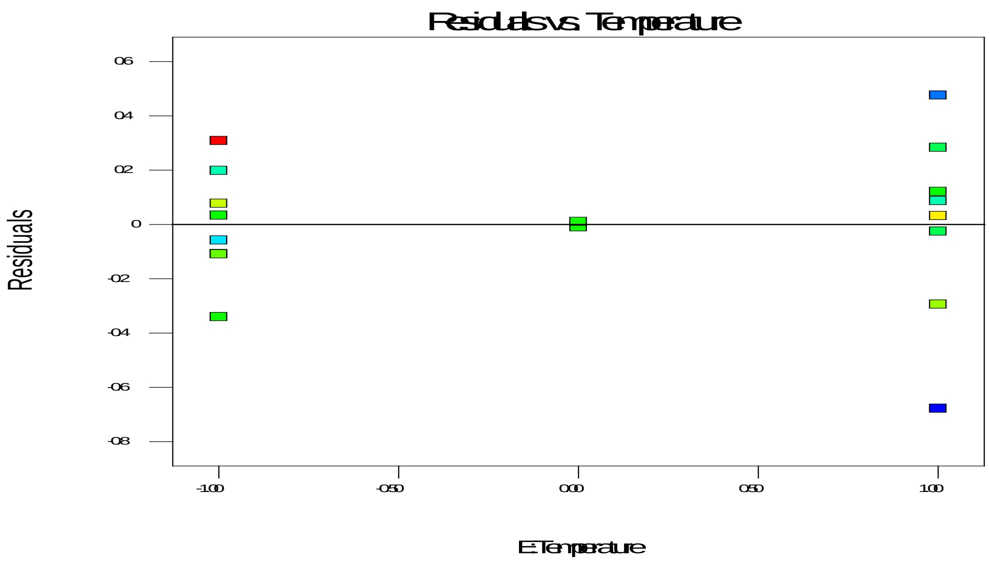

Residuals vs. Run
Residuals
Run Number


Residuals vs. SiH2Cl2
Residuals
SiH2Cl2

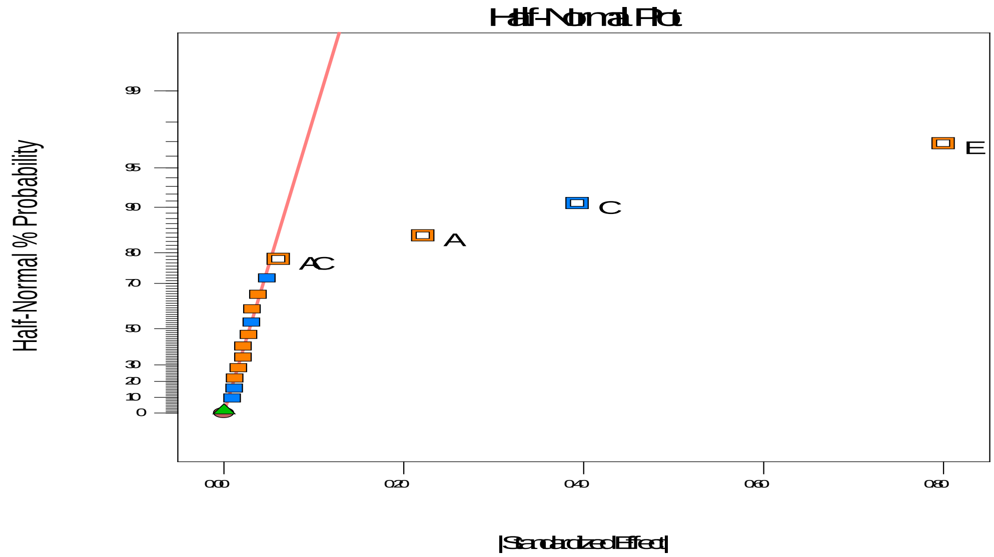

Residuals vs. C6H14
Residuals
B:C6H14


Residuals vs. HCl
Residuals
C:HCl





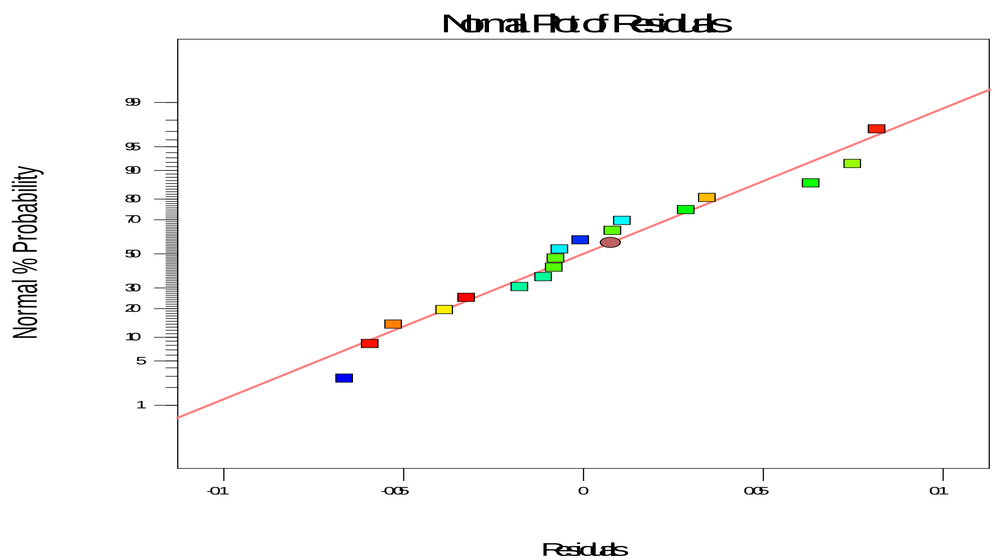
The following analysis is for the Si cap thickness response with a natural log transformation applied. The *A*, *C*, and *E* factors are still significant; however, the CE interaction is no longer significant. The residual plots are acceptable. The *AC* interaction is now significant. The curvature remains not significant.

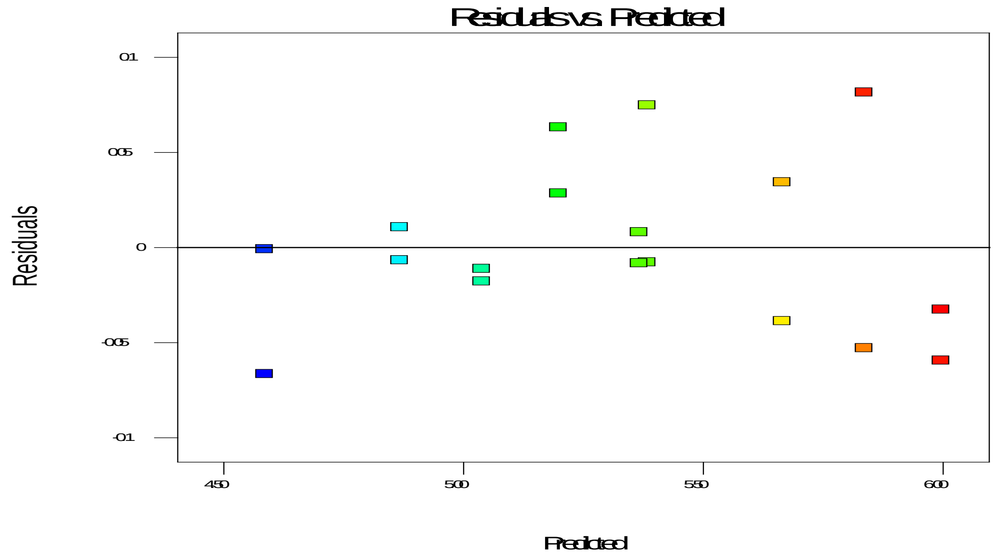


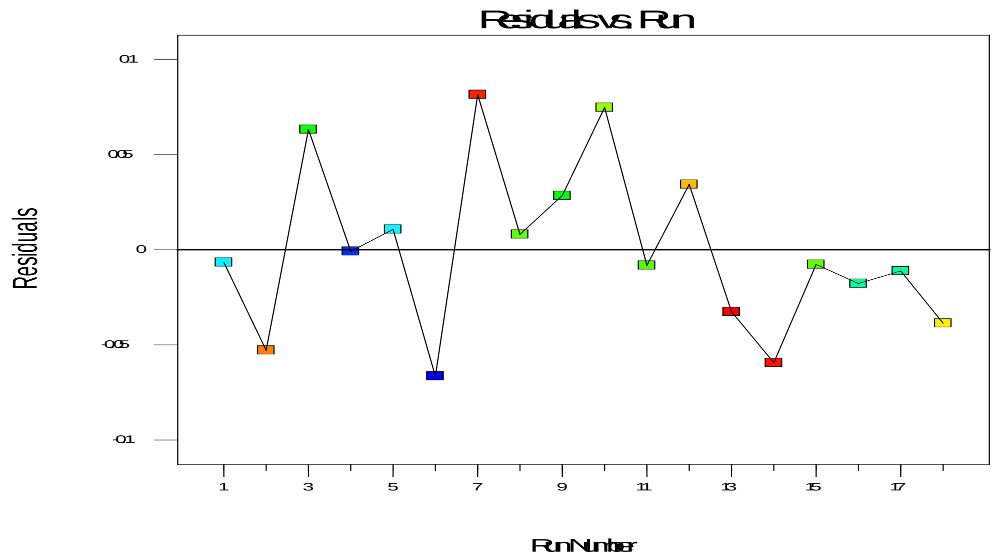
Design Expert Output

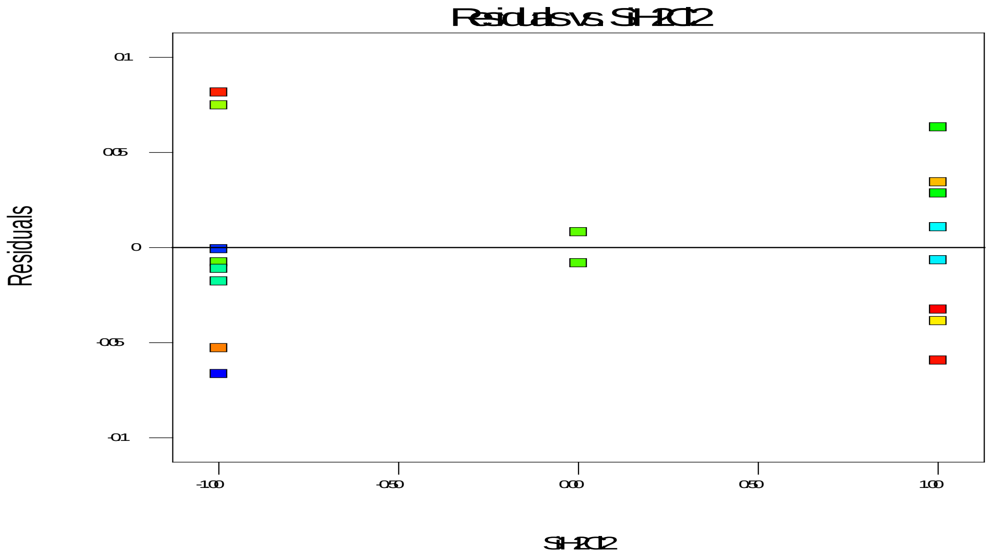
Response 1 Si cap thickness
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

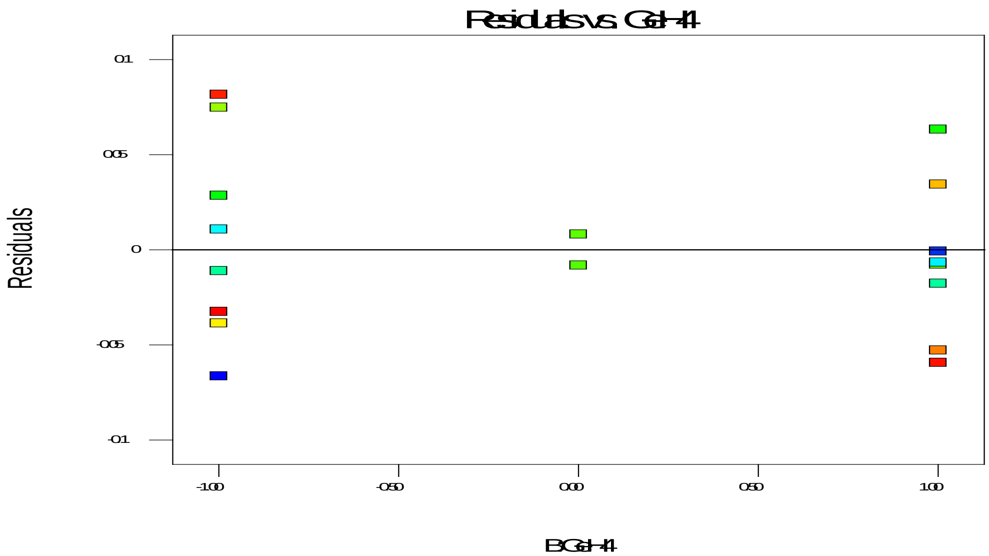
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	3.37	4	0.84	312.76	< 0.0001	significant
A-SiH2Cl2	*0.20*	*1*	*0.20*	*72.45*	*< 0.0001*	
C-HCl	*0.61*	*1*	*0.61*	*228.10*	*< 0.0001*	
E-Temperature	*2.55*	*1*	*2.55*	*945.03*	*< 0.0001*	
AC	*0.015*	*1*	*0.015*	*5.46*	*0.0376*	
Curvature	3.657E-003	1	3.657E-003	1.36	0.2665	not significant
Residual	0.032	12	2.693E-003			
Lack of Fit	*0.032*	*11*	*2.926E-003*	*21.99*	*0.1650*	*not significant*
Pure Error	*1.330E-004*	*1*	*1.330E-004*			
Cor Total	3.41	17				

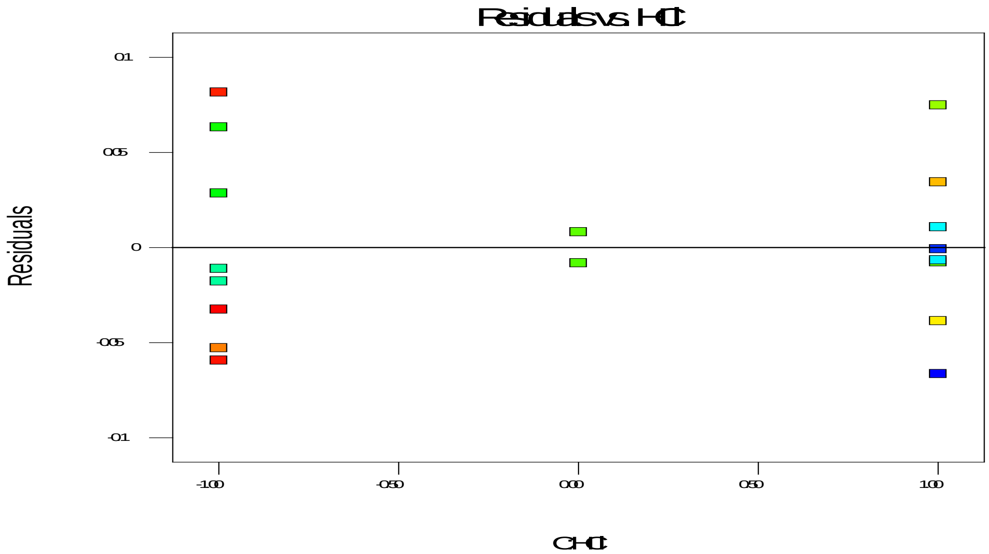


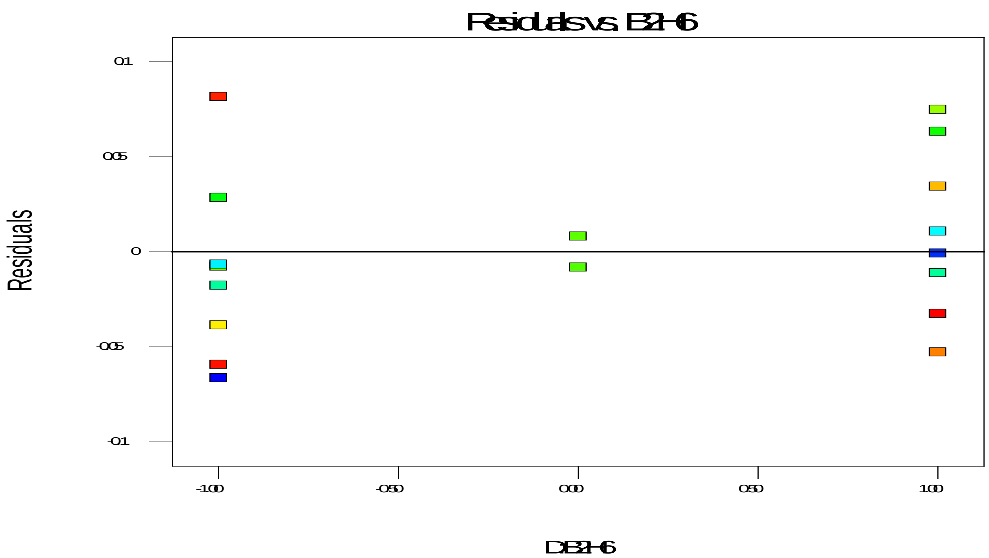

Residuals vs. Predicted
Residuals
Predicted
0.1
0.05
0
-0.05
-0.1
450
500
550
600

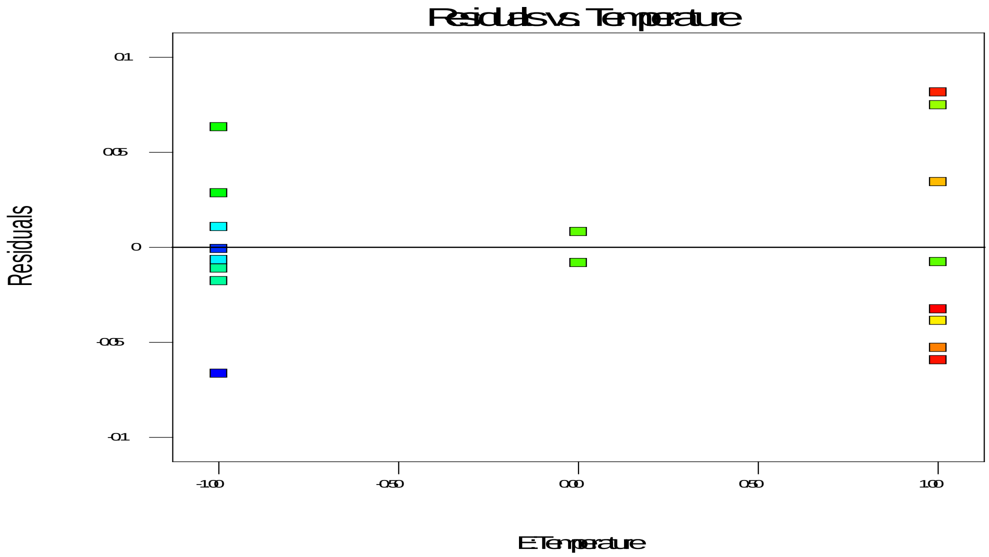

Residuals vs. Run
Residuals
Run Number
0.1
0.05
0
-0.05
-0.1
1
3
5
7
9
11
13
15
17


Residuals


Residuals


Residuals


Residuals



8.37. An article in *Soldering & Surface Mount Technology* ("Characterization of Solder Paste Printing Process and Its Optimization," 1999, Vol.11, No. 3, pp. 23-26) describes the use of a 2^{8-3} fractional factorial experiment to study the effect of eight factors on two responses: percentage volume matching (PVM) – the ration of the actual printed solder paste volume to the designed volume; and non-conformities per unit (NPU) – the number of solder paste printing defects determined by visual inspection (20' magnification) after printing according to an industry workmanship standard. The factor levels are shown below and the test matrix and response data are shown in Table 8.9.

	Levels	
Parameters	(-)	(+)
A. Squeegee pressure, *MPa*	0.1	0.3
B. Printing speed, *mm/s*	24	32
C. Squeegee angle, *deg*	45	65
D. Temperature, *°C*	20	28
E. Viscosity, *kCps*	1,100-1,150	1,250-1,300
F. Cleaning interval, *stroke*	8	15
G. Separation speed, *mm/s*	0.4	0.8
H. Relative humidity, *%*	30	70

Table P8.9 – The Solder Paste Experiment

Run	Parameters									
Order	A	B	C	D	E	F	G	H	PVM	NPU (%)
4	-	-	-	-	-	-	-	+	1.00	5
13	+	-	-	-	-	+	+	+	1.04	13
6	-	+	-	-	-	+	+	-	1.02	16
3	+	+	-	-	-	-	-	-	0.99	12
19	-	-	+	-	-	+	-	-	1.02	15
25	+	-	+	-	-	-	+	-	1.01	9
21	-	+	+	-	-	-	+	+	1.01	12
14	+	+	+	-	-	+	-	+	1.03	17
10	-	-	-	+	-	-	+	-	1.04	21

22	+	-	-	+	-	+	-	-	1.14	20
1	-	+	-	+	-	+	-	+	1.20	25
2	+	+	-	+	-	-	+	+	1.13	21
30	-	-	+	+	-	+	+	+	1.14	25
8	+	-	+	+	-	-	-	+	1.07	13
9	-	+	+	+	-	-	-	-	1.06	20
20	+	+	+	+	-	+	+	-	1.13	26
17	-	-	-	-	+	-	-	-	1.02	10
18	+	-	-	-	+	+	+	-	1.10	13
5	-	+	-	-	+	+	+	+	1.09	17
26	+	+	-	-	+	-	-	+	0.96	13
31	-	-	+	-	+	+	-	+	1.02	14
11	+	-	+	-	+	-	+	+	1.07	11
29	-	+	+	-	+	-	+	-	0.98	10
23	+	+	+	-	+	+	-	-	0.95	14
32	-	-	-	+	+	-	+	+	1.10	28
7	+	-	-	+	+	+	-	+	1.12	24
15	-	+	-	+	+	+	-	-	1.19	22
27	+	+	-	+	+	-	+	-	1.13	15
12	-	-	+	+	+	+	+	-	1.20	21
28	+	-	+	+	+	-	-	-	1.07	19
24	-	+	+	+	+	-	-	+	1.12	21
16	+	+	+	+	+	+	+	+	1.21	27

(a) Verify that the generators are $I = ABCF$, $I = ABDG$, and $I = BCDEH$ for this design.

This can be done by recognizing that Column $F = ABC$, Column $G = ABD$, and Column $H = BCDE$.

(b) What are the aliases for the main effects and two-factor interactions? You can ignore all interactions order three and higher.

The full defining relationship is: I + ABCF + ABDG + CDFG + ADEFH + ACEGH + BCDEH; and the alias structure for the main effects and two-factor interactions are:

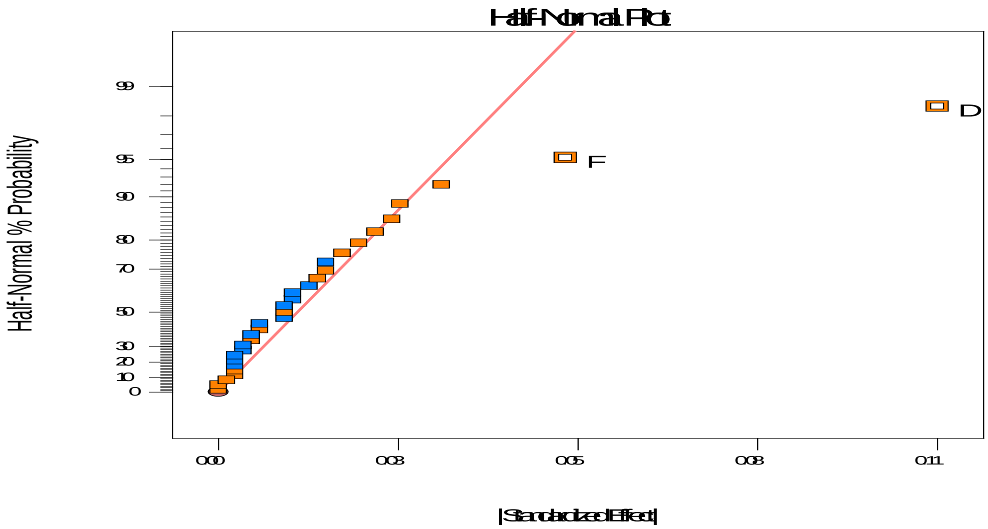
```
A + BCF + BDG + CEGH + DEFH
B + ACF + ADG + CDEH + EFGH
C + ABF + DFG + AEGH + BDEH
D + ABG + CFG + AEFH + BCEH
E + ACGH + ADFH + BCDH + BFGH
F + ABC + CDG + ADEH + BEGH
G + ABD + CDF + ACEH + BEFH
H + ACEG + ADEF + BCDE + BEFG
AB + CF + DG
AC + BF + EGH + ADFG + BCDG
AD + BG + EFH + ACFG + BCDF
AE + CGH + DFH + BCEF + BDEG
AF + BC + DEH + ACDG + BDFG
AG + BD + CEH + ACDF + BCFG
AH + CEG + DEF + BCFH + BDGH
BE + CDH + FGH + ACEF + ADEG
BH + CDE + EFG + ACFH + ADGH
CD + FG + BEH + ABCG + ABDF
CE + AGH + BDH + ABEF + DEFG
CG + DF + AEH + ABCD + ABFG
CH + AEG + BDE + ABFH + DFGH
DE + AFH + BCH + ABEG + CEFG
DH + AEF + BCE + ABGH + CFGH
EF + ADH + BGH + ABCE + CDEG
EG + ACH + BFH + ABDE + CDEF
```

```
EH + ACG + ADF + BCD + BFG
FH + ADE + BEG + ABCH + CDGH
GH + ACE + BEF + ABDH + CDFH
```

(c) Analyze both the PVM and NPU responses.

For both the PVM and NPU responses, only factors D and F appear to be significant.

The PVM analysis is shown below.



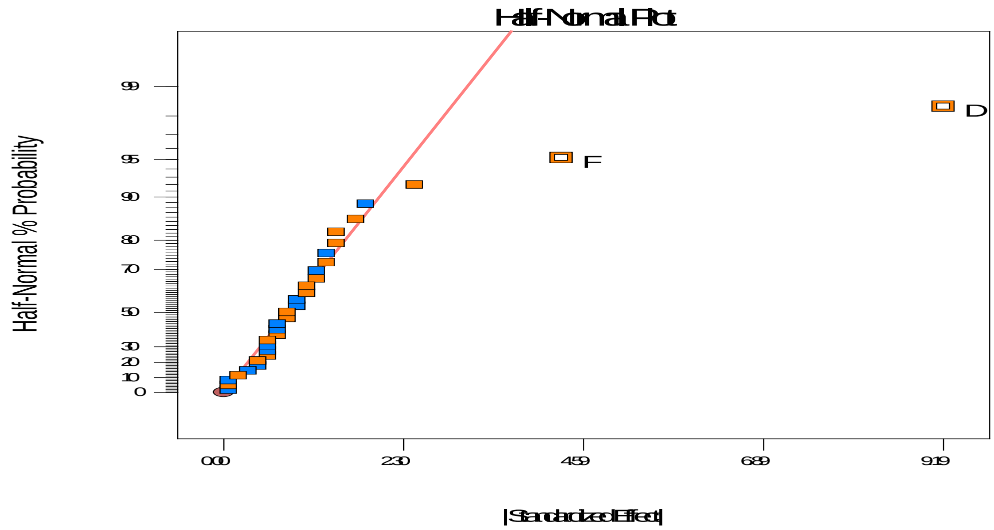
Design Expert Output

Response 1 PVM
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	0.12	2	0.058	37.52	< 0.0001	significant
D-Temperature	*0.095*	*1*	*0.095*	*60.85*	*< 0.0001*	
F-Cleaning interval	*0.022*	*1*	*0.022*	*14.18*	*0.0008*	
Residual	0.045	29	1.555E-003			
Cor Total	0.16	31				

Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	1.07	1	6.970E-003	1.06	1.09	
D-Temperature	0.054	1	6.970E-003	0.040	0.069	1.00
F-Cleaning interval	0.026	1	6.970E-003	0.012	0.041	1.00

The analysis for the NPU response is shown below.



Design Expert Output

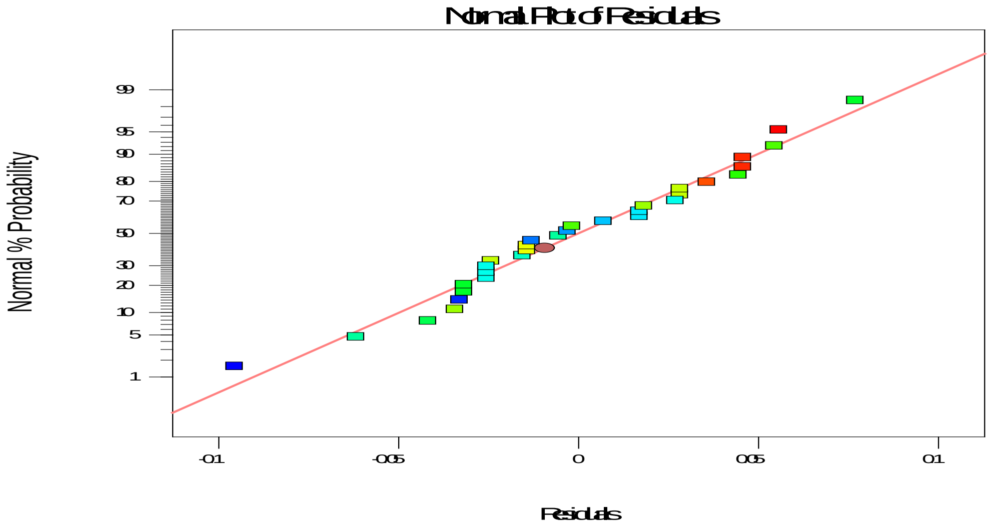
Response 2 NPU
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

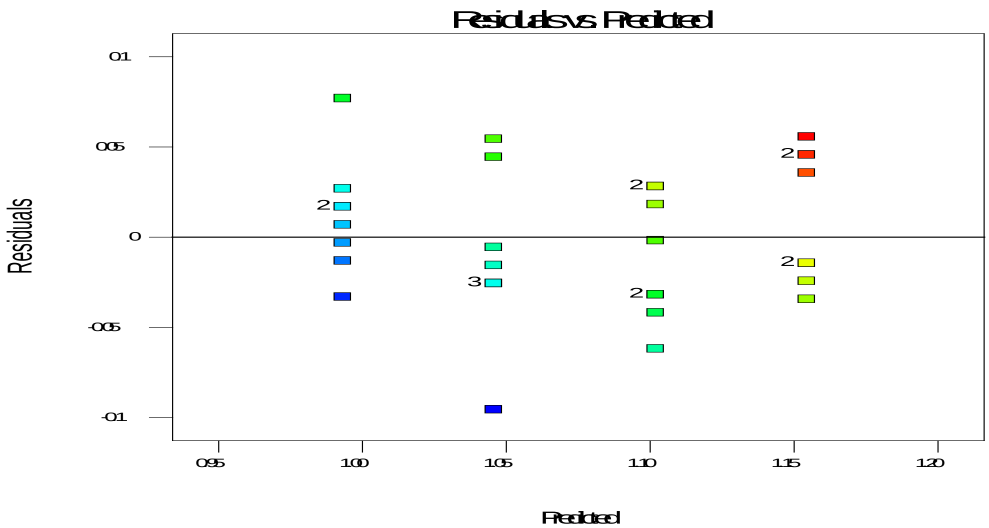
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	824.06	2	412.03	48.15	< 0.0001	significant
D-Temperature	*675.28*	*1*	*675.28*	*78.91*	*< 0.0001*	
F-Cleaning interval	*148.78*	*1*	*148.78*	*17.39*	*0.0003*	
Residual	248.16	29	8.56			
Cor Total	1072.22	31				

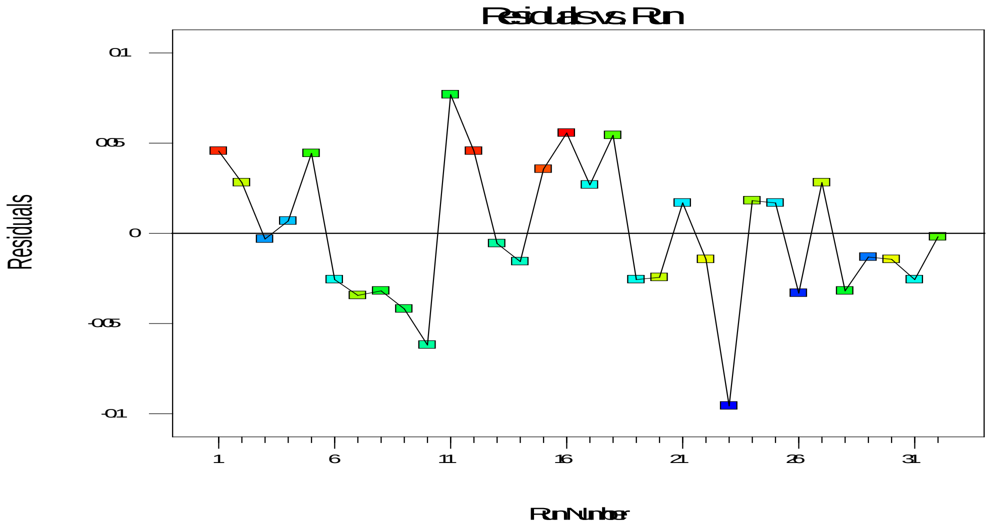
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	17.16	1	0.52	16.10	18.21	
D-Temperature	4.59	1	0.52	3.54	5.65	1.00
F-Cleaning interval	2.16	1	0.52	1.10	3.21	1.00

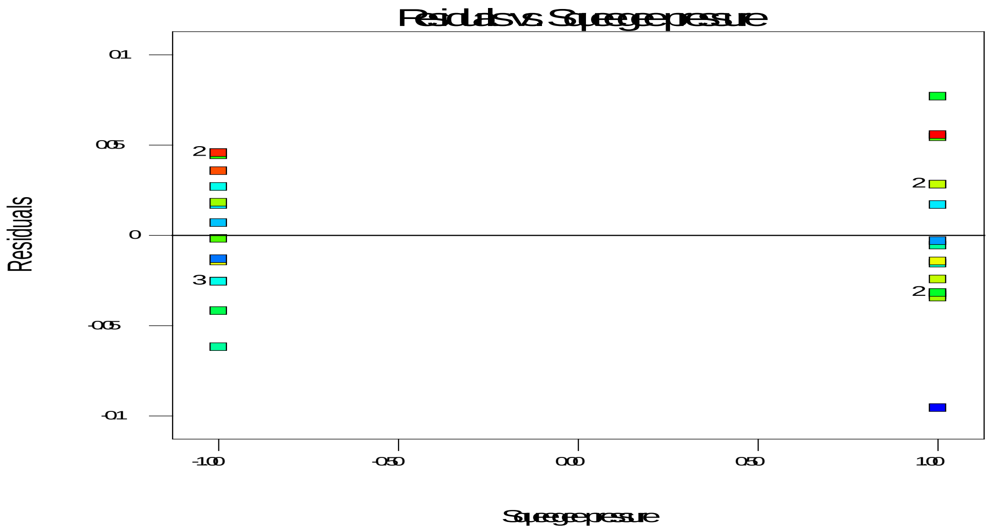
(d) Analyze the residuals for both responses. Are there any problems with model adequacy?

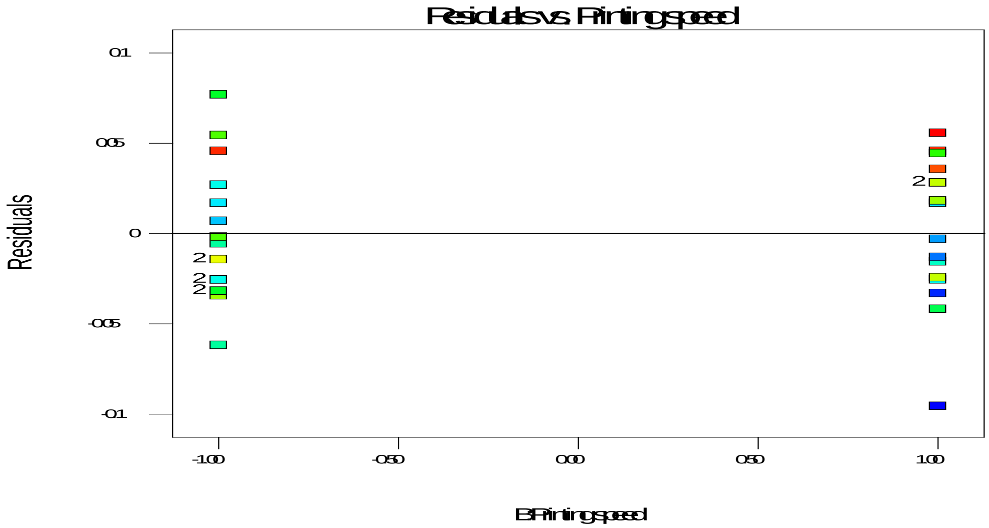
The residual plots for the PVM response shown below do not identify any concerns.

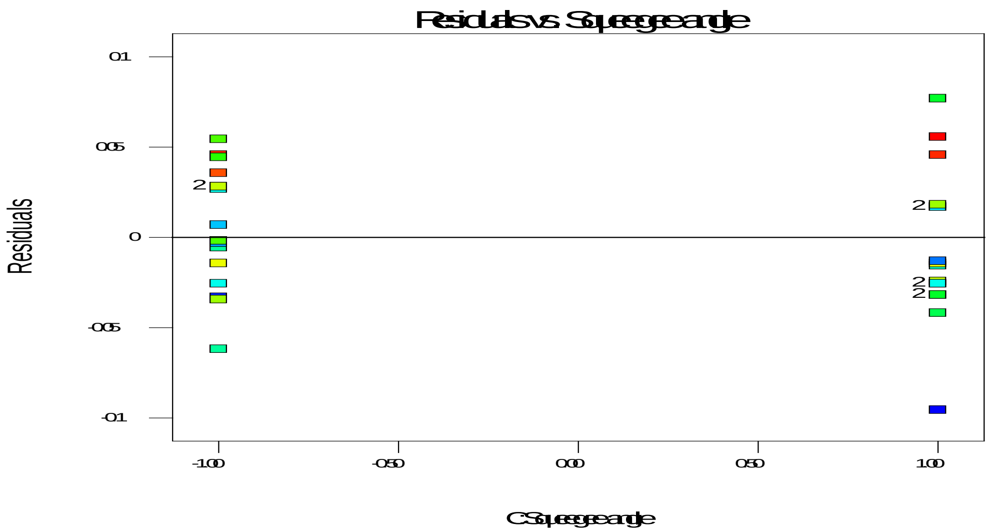

Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
Residuals

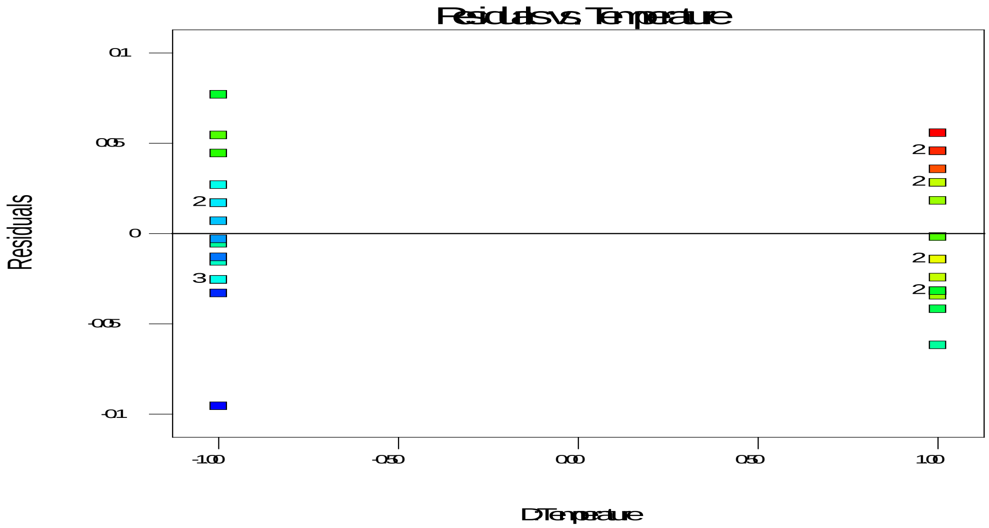

Residuals vs. Predicted
Residuals
0
Predicted

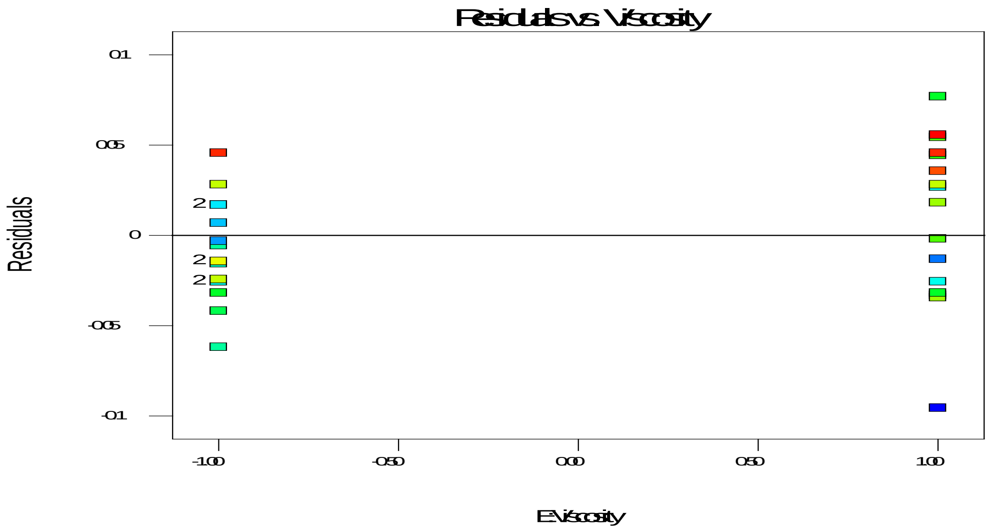

Residuals vs. Run
Residuals
0.1
0.05
0
-0.05
-0.1
1
6
11
16
21
26
31
Run Number

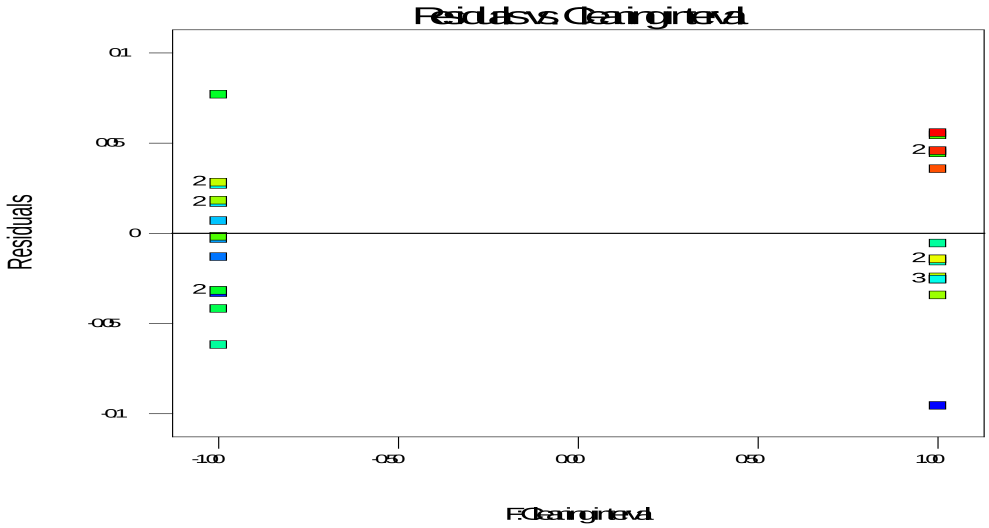

Residuals vs. Squeegee pressure
Residuals
0.1
0.05
0
-0.05
-0.1
2
3
2
2
-1.00
-0.50
0.00
0.50
1.00
Squeegee pressure

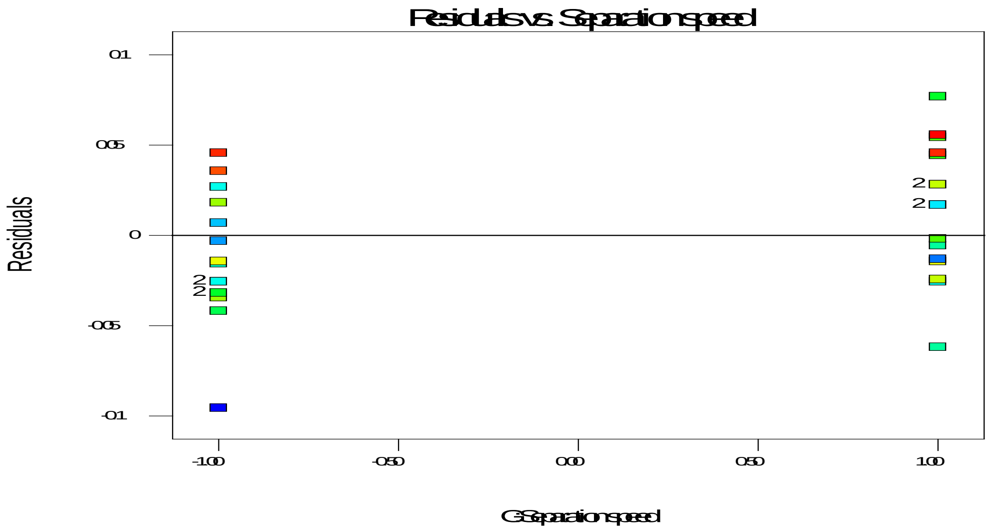

Residuals vs. Printing speed
Residuals
0.1
0.05
0
-0.05
-0.1
-1.00
-0.50
0.00
0.50
1.00
B:Printing speed

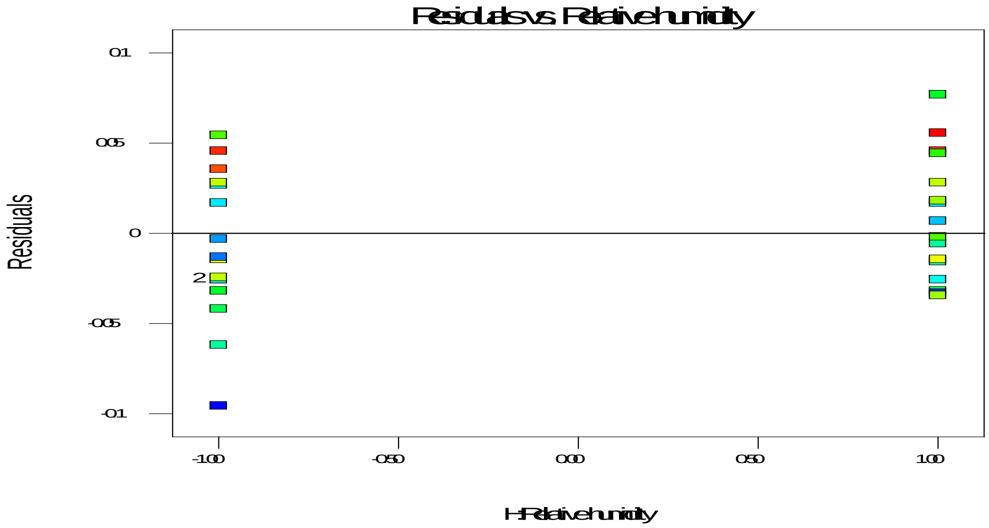

Residuals vs. Squeegee angle
Residuals
0.1
0.05
0
-0.05
-0.1
-1.00
-0.50
0.00
0.50
1.00
C:Squeegee angle


Residuals vs. Temperature
Residuals
0.1
0.05
0
-0.05
-0.1
-1.00
-0.50
0.00
0.50
1.00
D:Temperature

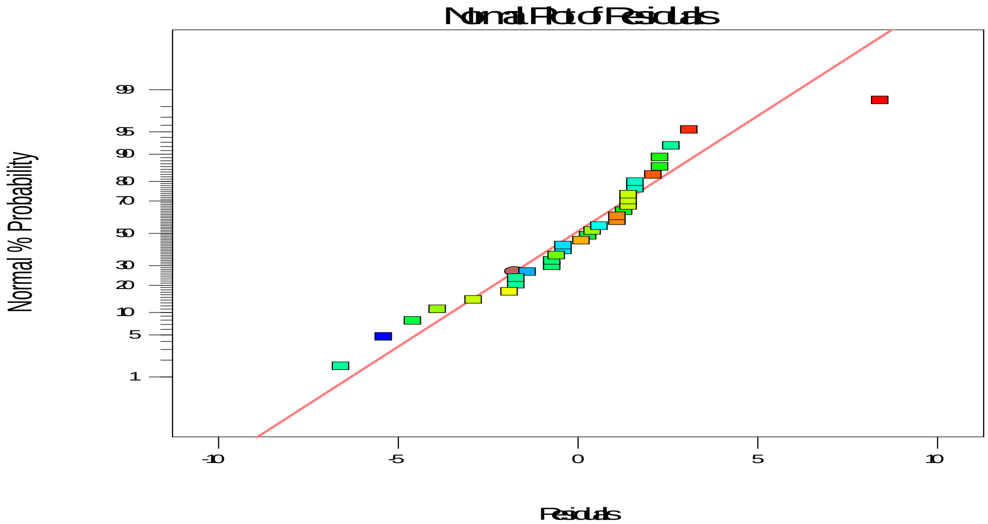

Residuals vs. Viscosity
Residuals
0.1
0.05
0
-0.05
-0.1
-1.00
-0.50
0.00
0.50
1.00
E:Viscosity

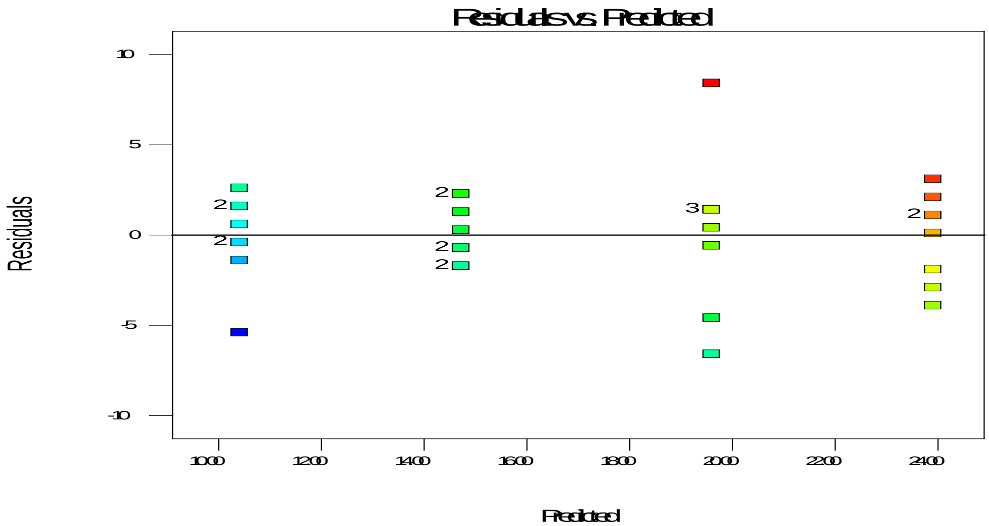

Residuals vs. Cleaning interval
Residuals
0.1
0.05
0
-0.05
-0.1
-1.00
-0.50
0.00
0.50
1.00
F: Cleaning interval

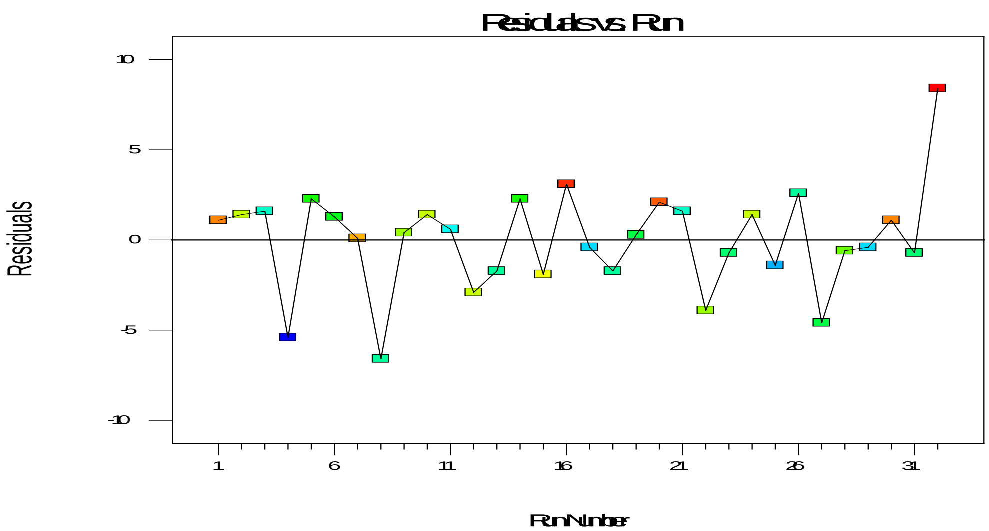

Residuals vs. Separation speed
Residuals
0.1
0.05
0
-0.05
-0.1
-1.00
-0.50
0.00
0.50
1.00
G: Separation speed

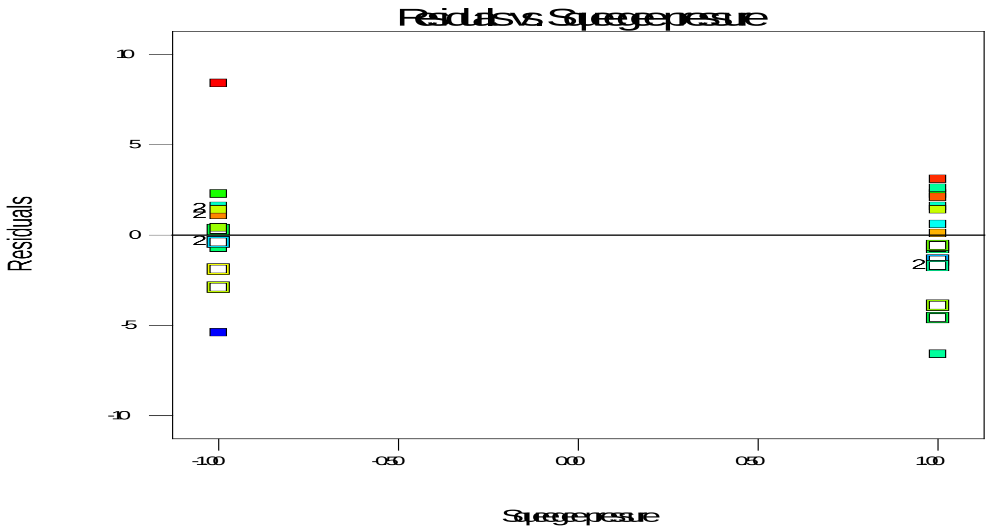


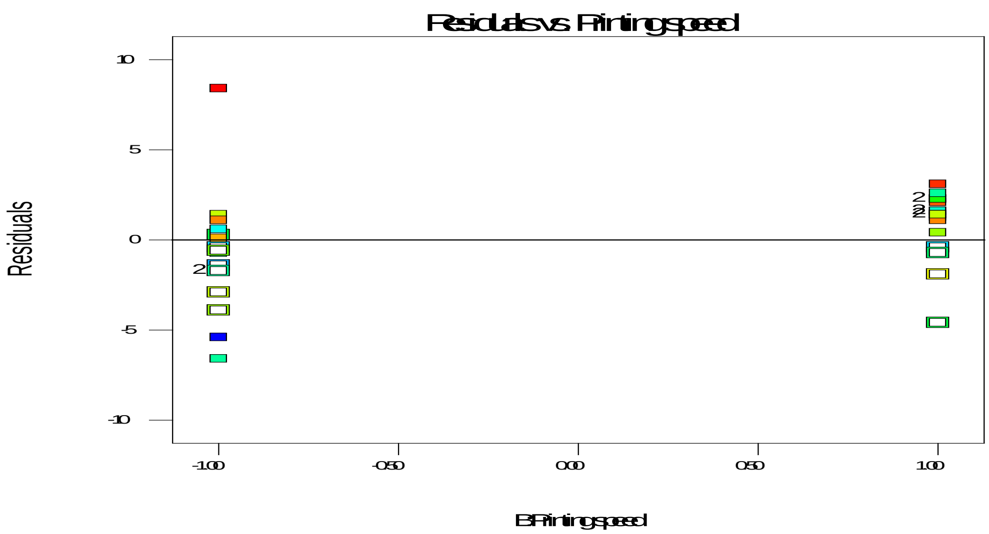
The residual plots for the NPU response shown below identify an outlier with run number 32, NPU value of 28. The experimenters should investigate this run.

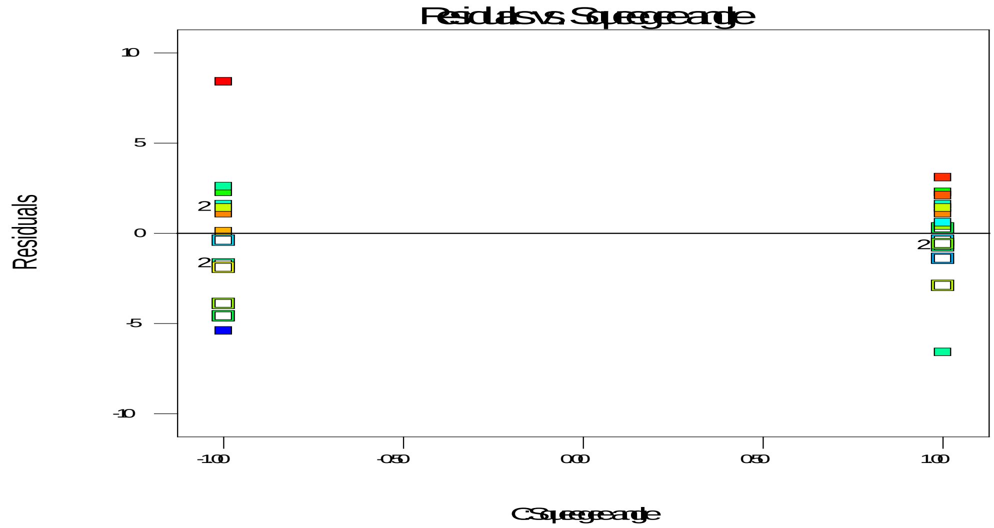


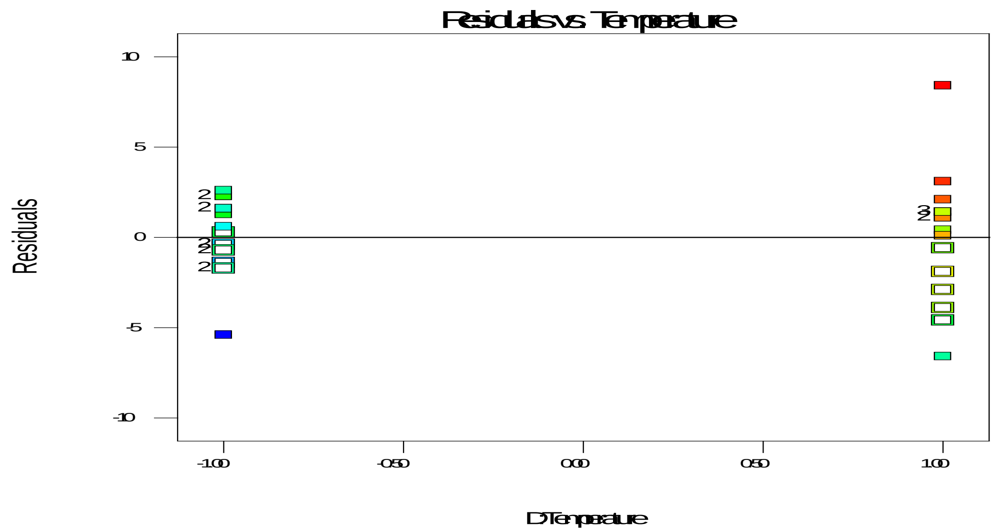

Residuals vs. Predicted
Residuals
10
5
0
-5
-10
2
2
2
2
2
3
2
100
120
140
160
180
200
220
240
Predicted

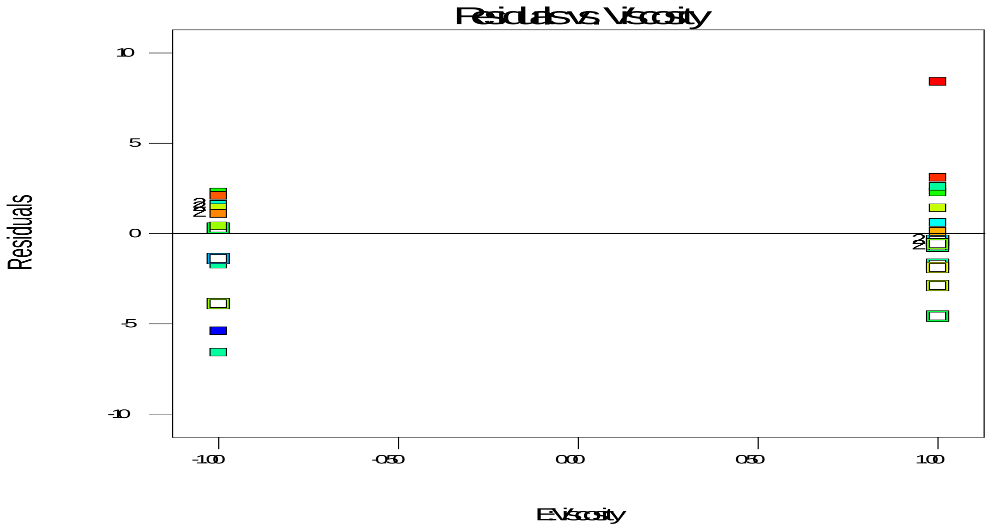

Residuals vs. Run
Residuals
10
5
0
-5
-10
1
6
11
16
21
26
31
Run Number

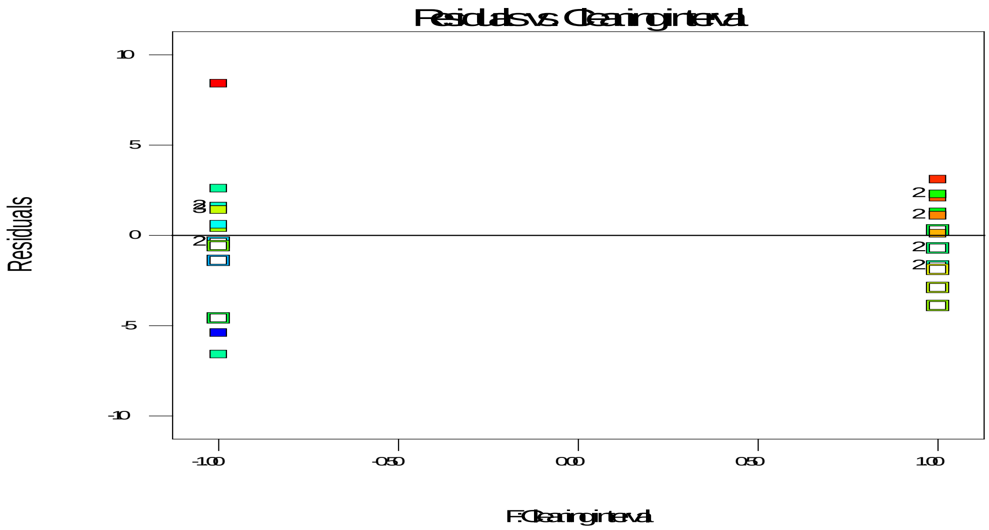

Residuals vs. Squeegee pressure
Residuals
10
5
0
-5
-10
-1.00
-0.50
0.00
0.50
1.00
Squeegee pressure

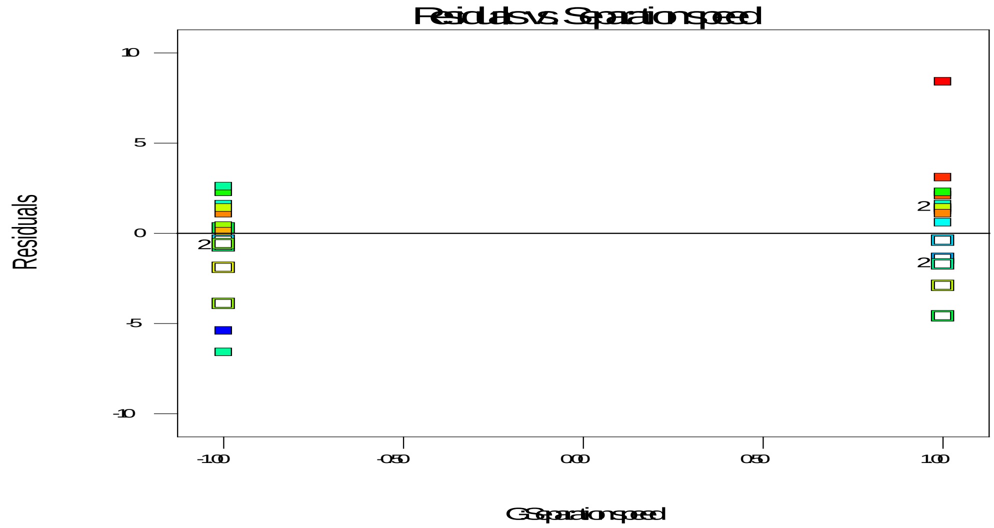

Residuals vs. Printing speed
Residuals
10
5
0
-5
-10
-1.00
-0.50
0.00
0.50
1.00
B:Printing speed

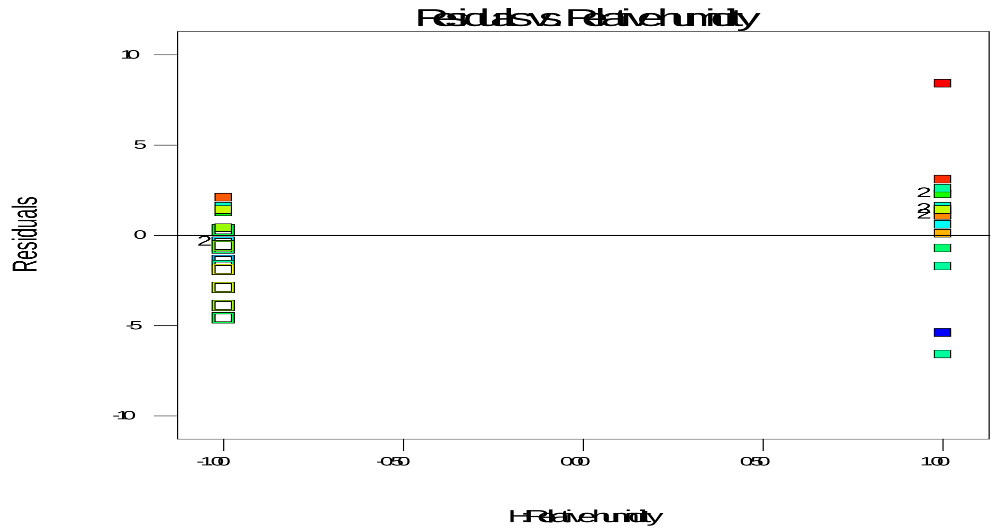

Residuals vs Squeegee angle
Residuals
10
5
0
-5
-10
-1.00
-0.50
0.00
0.50
1.00
C:Squeegee angle


Residuals vs Temperature
Residuals
10
5
0
-5
-10
-1.00
-0.50
0.00
0.50
1.00
D:Temperature


Residuals vs. Viscosity
Residuals
10
5
0
-5
-10
-1.00
-0.50
0.00
0.50
1.00
E:Viscosity

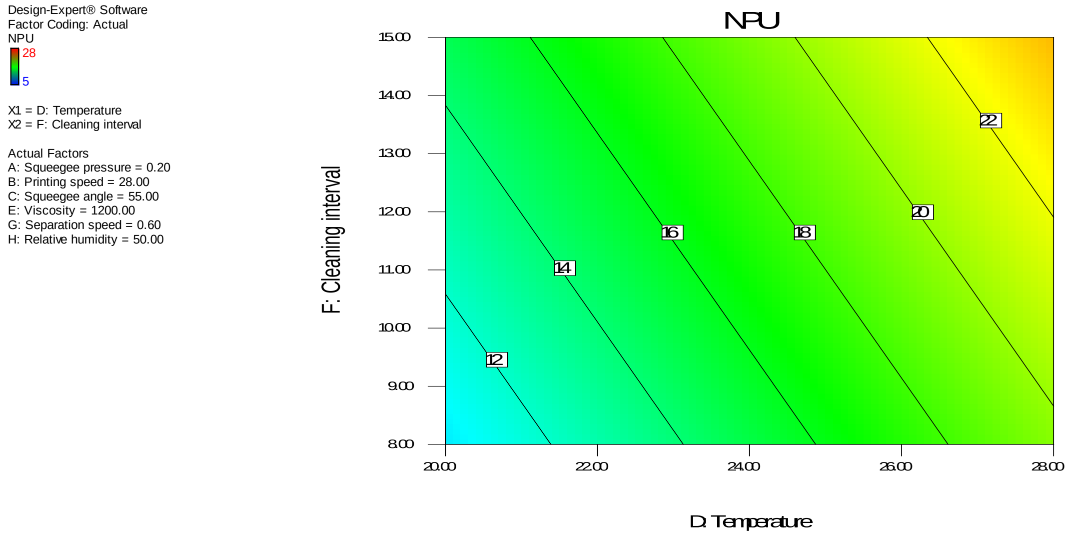

Residuals vs. Cleaning interval
Residuals
10
5
0
-5
-10
-1.00
-0.50
0.00
0.50
1.00
F:Cleaning interval

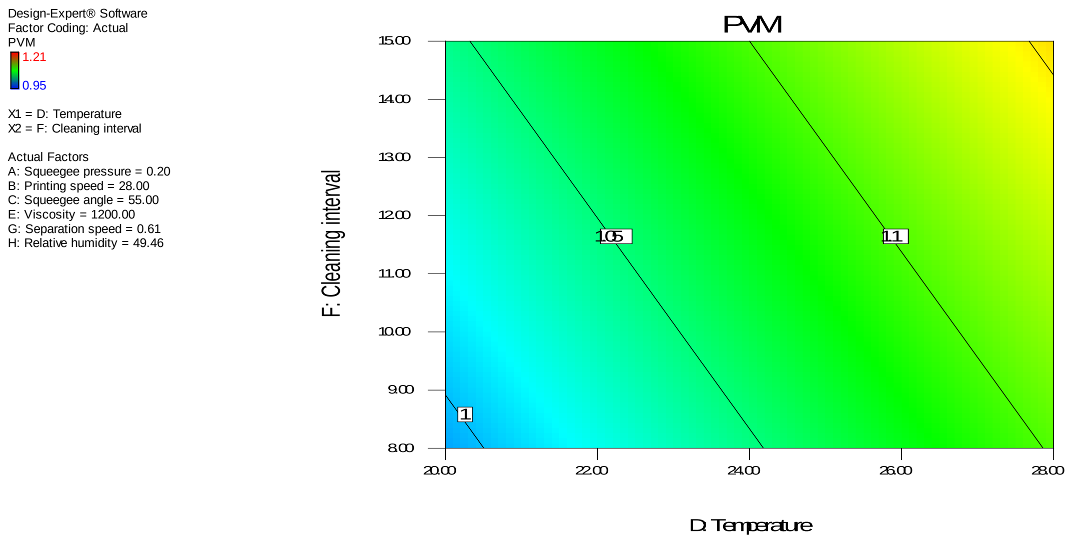


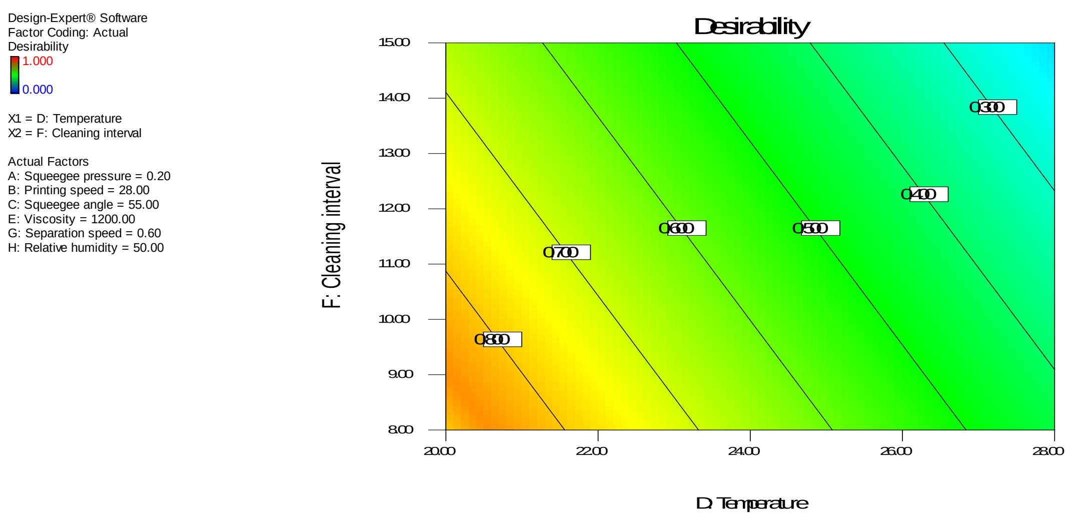


(e) The ideal value of PVM is unity and the NPU response should be as small as possible. Recommend suitable operating conditions for the process based on the experimental results.

The contour plots of both the PVM and NPU responses show that running the Temperature at 20C and the Cleaning Interval at 8 is the best operating conditions. This is confirmed with the desirability plot also shown below. The other factors should be set based on process controls, costs, robustness, or other process performance requirements, as they are not significant to either the PVM or NPU response.


Design-Expert® Software
Factor Coding: Actual
NPU
28
5
X1 = D: Temperature
X2 = F: Cleaning interval
Actual Factors
A: Squeegee pressure = 0.20
B: Printing speed = 28.00
C: Squeegee angle = 55.00
E: Viscosity = 1200.00
G: Separation speed = 0.60
H: Relative humidity = 50.00
NPU
F: Cleaning interval
D: Temperature
15.00
14.00
13.00
12.00
11.00
10.00
9.00
8.00
20.00
22.00
24.00
26.00
28.00
12
14
16
18
20
22


Design-Expert® Software
Factor Coding: Actual
PVM
1.21
0.95
X1 = D: Temperature
X2 = F: Cleaning interval
Actual Factors
A: Squeegee pressure = 0.20
B: Printing speed = 28.00
C: Squeegee angle = 55.00
E: Viscosity = 1200.00
G: Separation speed = 0.61
H: Relative humidity = 49.46
PVM
F: Cleaning interval
D: Temperature
15.00
14.00
13.00
12.00
11.00
10.00
9.00
8.00
20.00
22.00
24.00
26.00
28.00
1
1.05
1.1


Design-Expert® Software
Factor Coding: Actual
Desirability
1.000
0.000
X1 = D: Temperature
X2 = F: Cleaning interval
Actual Factors
A: Squeegee pressure = 0.20
B: Printing speed = 28.00
C: Squeegee angle = 55.00
E: Viscosity = 1200.00
G: Separation speed = 0.60
H: Relative humidity = 50.00
Desirability
F: Cleaning interval
D: Temperature
15.00
14.00
13.00
12.00
11.00
10.00
9.00
8.00
20.00
22.00
24.00
26.00
28.00
0.800
0.700
0.600
0.500
0.400
0.300

8.38. An article in the *International Journal of Research in Marketing* ("Experimental design on the front lines of marketing: Testing new ideas to increase direct mail sales," 2006. Vol. 23, pp. 309-319) describes the use of a 20-run Plackett-Burman design to improve the effects of 19 factors to improve the response rate to a direct mail campaign to attract new customers to a credit card. The 19 factors are as follows:

Factor	(-) Control	(+) New Idea
A: Envelope teaser	General offer	Product-specific offer
B: Return address	Blind	Add company name
C: "Official" ink-stamp on envelope	Yes	No
D: Postage	Pre-printed	Stamp
E: Additional graphic on envelop	Yes	No
F: Price graphic on letter	Small	Large
G: Sticker	Yes	No
H: Personalize letter copy	No	Yes
I: Copy Message	Targeted	Generic
J: Letter headline	Headline 1	Headline 2
K: List of Benefits	Standard layout	Creative layout
L: Postscript on letter	Control version	New P.S.
M: Signature	Manager	Senior Executive
N: Product selection	Many	Few
O: Value of free gift	High	Low
P: Reply envelope	Control	New Style
Q: Information on buckslip	Product info	Free gift info
R: 2nd buckslip	No	Yes
S: Interest rate	Low	High

The 20-run Plackett-Burman design is shown in the following table. Each test combination in the table was mailed to 5,000 potential customers, and the response rate is the percentage of customers who responded positively to the offer.

Test Cell	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	Orders	Resp Rate
1	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	52	1.04%
2	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	38	0.76%
3	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	42	0.84%
4	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	134	2.68%
5	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	104	2.08%
6	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	60	1.20%
7	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	61	1.22%
8	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	68	1.36%
9	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	57	1.14%
10	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	30	0.60%
11	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	108	2.16%
12	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	39	0.78%
13	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	40	0.80%
14	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	49	0.98%
15	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	37	0.74%
16	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	99	1.98%

17	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	86	1.72%
18	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	43	0.86%
19	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	47	0.94%
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	104	2.08%

(a) Verify that in this design each main effect is aliased to all two-factor interactions except those that involve that main effect. That is, in the 19 factor design, the main effect for each factor is aliased with all two-factor interactions involving the other 18 factors, or 153 two-factor interactions (18!/2!16!).

The Design Expert output shown below identifies the aliases for the main effects and the two factor interactions. Notice that the main effects for each factor is aliased with all two-factor interactions involving the 18 other factors, but do not include the interactions that involve the same main effect.

Design Expert Output

```
[A] = A - 0.2 * BC + 0.2 * BD + 0.2 * BE + 0.2 * BF - 0.2 * BG + 0.2 * BH
 - 0.6 * BI - 0.2 * BJ - 0.2 * BK + 0.2 * BL + 0.2 * BM - 0.2 * BN + 0.2 * BO
 - 0.2 * BP - 0.2 * BQ - 0.2 * BR - 0.2 * BS - 0.2 * CD - 0.2 * CE - 0.6 * CF
 - 0.2 * CG + 0.2 * CH - 0.2 * CI - 0.2 * CJ - 0.2 * CK + 0.2 * CL + 0.2 * CM
 + 0.2 * CN - 0.2 * CO + 0.2 * CP + 0.2 * CQ - 0.2 * CR + 0.2 * CS - 0.2 * DE
 + 0.2 * DF - 0.2 * DG + 0.2 * DH + 0.2 * DI - 0.2 * DJ + 0.2 * DK - 0.2 * DL
 - 0.2 * DM + 0.2 * DN - 0.2 * DO - 0.2 * DP - 0.2 * DQ - 0.6 * DR + 0.2 * DS
 - 0.2 * EF + 0.2 * EG - 0.2 * EH - 0.2 * EI + 0.2 * EJ + 0.2 * EK - 0.2 * EL
 + 0.2 * EM - 0.6 * EN - 0.2 * EO - 0.2 * EP + 0.2 * EQ - 0.2 * ER + 0.2 * ES
 + 0.2 * FG - 0.2 * FH - 0.2 * FI - 0.2 * FJ - 0.2 * FK - 0.2 * FL - 0.2 * FM
 + 0.2 * FN - 0.2 * FO + 0.2 * FP + 0.2 * FQ + 0.2 * FR - 0.2 * FS - 0.2 * GH
 + 0.2 * GI + 0.2 * GJ - 0.6 * GK + 0.2 * GL - 0.2 * GM - 0.2 * GN - 0.2 * GO
 + 0.2 * GP - 0.2 * GQ - 0.2 * GR + 0.2 * GS + 0.2 * HI + 0.2 * HJ - 0.2 * HK
 + 0.2 * HL - 0.2 * HM - 0.2 * HN - 0.2 * HO - 0.2 * HP + 0.2 * HQ - 0.2 * HR
 - 0.6 * HS + 0.2 * IJ + 0.2 * IK - 0.2 * IL - 0.2 * IM - 0.2 * IN + 0.2 * IO
 + 0.2 * IP - 0.2 * IQ - 0.2 * IR - 0.2 * IS - 0.2 * JK - 0.2 * JL + 0.2 * JM
 + 0.2 * JN - 0.2 * JO - 0.6 * JP - 0.2 * JQ + 0.2 * JR - 0.2 * JS - 0.2 * KL
 - 0.2 * KM - 0.2 * KN + 0.2 * KO - 0.2 * KP + 0.2 * KQ + 0.2 * KR + 0.2 * KS
 - 0.6 * LM - 0.2 * LN + 0.2 * LO - 0.2 * LP - 0.2 * LQ + 0.2 * LR + 0.2 * LS
 + 0.2 * MN + 0.2 * MO + 0.2 * MP - 0.2 * MQ - 0.2 * MR - 0.2 * MS - 0.2 * NO
 - 0.2 * NP - 0.2 * NQ + 0.2 * NR + 0.2 * NS + 0.2 * OP - 0.6 * OQ + 0.2 * OR
 - 0.2 * OS + 0.2 * PQ - 0.2 * PR - 0.2 * PS + 0.2 * QR - 0.2 * QS - 0.2 * RS
[B] = B - 0.2 * AC + 0.2 * AD + 0.2 * AE + 0.2 * AF - 0.2 * AG + 0.2 * AH
 - 0.6 * AI - 0.2 * AJ - 0.2 * AK + 0.2 * AL + 0.2 * AM - 0.2 * AN + 0.2 * AO
 - 0.2 * AP - 0.2 * AQ - 0.2 * AR - 0.2 * AS - 0.2 * CD + 0.2 * CE + 0.2 * CF
 + 0.2 * CG - 0.2 * CH + 0.2 * CI - 0.6 * CJ - 0.2 * CK - 0.2 * CL + 0.2 * CM
 + 0.2 * CN - 0.2 * CO + 0.2 * CP - 0.2 * CQ - 0.2 * CR - 0.2 * CS - 0.2 * DE
 - 0.2 * DF - 0.6 * DG - 0.2 * DH + 0.2 * DI - 0.2 * DJ - 0.2 * DK - 0.2 * DL
 + 0.2 * DM + 0.2 * DN + 0.2 * DO - 0.2 * DP + 0.2 * DQ + 0.2 * DR - 0.2 * DS
 - 0.2 * EF + 0.2 * EG - 0.2 * EH + 0.2 * EI + 0.2 * EJ - 0.2 * EK + 0.2 * EL
 - 0.2 * EM - 0.2 * EN + 0.2 * EO - 0.2 * EP - 0.2 * EQ - 0.2 * ER - 0.6 * ES
 - 0.2 * FG + 0.2 * FH - 0.2 * FI - 0.2 * FJ + 0.2 * FK + 0.2 * FL - 0.2 * FM
 + 0.2 * FN - 0.6 * FO - 0.2 * FP - 0.2 * FQ + 0.2 * FR - 0.2 * FS + 0.2 * GH
 - 0.2 * GI - 0.2 * GJ - 0.2 * GK - 0.2 * GL - 0.2 * GM - 0.2 * GN + 0.2 * GO
 - 0.2 * GP + 0.2 * GQ + 0.2 * GR + 0.2 * GS - 0.2 * HI + 0.2 * HJ + 0.2 * HK
 - 0.6 * HL + 0.2 * HM - 0.2 * HN - 0.2 * HO - 0.2 * HP + 0.2 * HQ - 0.2 * HR
 - 0.2 * HS + 0.2 * IJ + 0.2 * IK - 0.2 * IL + 0.2 * IM - 0.2 * IN - 0.2 * IO
 - 0.2 * IP - 0.2 * IQ + 0.2 * IR - 0.2 * IS + 0.2 * JK + 0.2 * JL - 0.2 * JM
 - 0.2 * JN - 0.2 * JO + 0.2 * JP + 0.2 * JQ - 0.2 * JR - 0.2 * JS - 0.2 * KL
 - 0.2 * KM + 0.2 * KN + 0.2 * KO - 0.2 * KP - 0.6 * KQ - 0.2 * KR + 0.2 * KS
 - 0.2 * LM - 0.2 * LN - 0.2 * LO + 0.2 * LP - 0.2 * LQ + 0.2 * LR + 0.2 * LS
 - 0.6 * MN - 0.2 * MO + 0.2 * MP - 0.2 * MQ - 0.2 * MR + 0.2 * MS + 0.2 * NO
 + 0.2 * NP + 0.2 * NQ - 0.2 * NR - 0.2 * NS - 0.2 * OP - 0.2 * OQ - 0.2 * OR
 + 0.2 * OS + 0.2 * PQ - 0.6 * PR + 0.2 * PS + 0.2 * QR - 0.2 * QS + 0.2 * RS
[C] = C - 0.2 * AB - 0.2 * AD - 0.2 * AE - 0.6 * AF - 0.2 * AG + 0.2 * AH
 - 0.2 * AI - 0.2 * AJ - 0.2 * AK + 0.2 * AL + 0.2 * AM + 0.2 * AN - 0.2 * AO
 + 0.2 * AP + 0.2 * AQ - 0.2 * AR + 0.2 * AS - 0.2 * BD + 0.2 * BE + 0.2 * BF
 + 0.2 * BG - 0.2 * BH + 0.2 * BI - 0.6 * BJ - 0.2 * BK - 0.2 * BL + 0.2 * BM
 + 0.2 * BN - 0.2 * BO + 0.2 * BP - 0.2 * BQ - 0.2 * BR - 0.2 * BS - 0.2 * DE
 + 0.2 * DF + 0.2 * DG + 0.2 * DH - 0.2 * DI + 0.2 * DJ - 0.6 * DK - 0.2 * DL
 - 0.2 * DM + 0.2 * DN + 0.2 * DO - 0.2 * DP + 0.2 * DQ - 0.2 * DR - 0.2 * DS
 - 0.2 * EF - 0.2 * EG - 0.6 * EH - 0.2 * EI + 0.2 * EJ - 0.2 * EK - 0.2 * EL
```

```
 - 0.2 * EM + 0.2 * EN + 0.2 * EO + 0.2 * EP - 0.2 * EQ + 0.2 * ER + 0.2 * ES
 - 0.2 * FG + 0.2 * FH - 0.2 * FI + 0.2 * FJ + 0.2 * FK - 0.2 * FL + 0.2 * FM
 - 0.2 * FN - 0.2 * FO + 0.2 * FP - 0.2 * FQ - 0.2 * FR - 0.2 * FS - 0.2 * GH
 + 0.2 * GI - 0.2 * GJ - 0.2 * GK + 0.2 * GL + 0.2 * GM - 0.2 * GN + 0.2 * GO
 - 0.6 * GP - 0.2 * GQ - 0.2 * GR + 0.2 * GS + 0.2 * HI - 0.2 * HJ - 0.2 * HK
 - 0.2 * HL - 0.2 * HM - 0.2 * HN - 0.2 * HO + 0.2 * HP - 0.2 * HQ + 0.2 * HR
 + 0.2 * HS - 0.2 * IJ + 0.2 * IK + 0.2 * IL - 0.6 * IM + 0.2 * IN - 0.2 * IO
 - 0.2 * IP - 0.2 * IQ + 0.2 * IR - 0.2 * IS + 0.2 * JK + 0.2 * JL - 0.2 * JM
 + 0.2 * JN - 0.2 * JO - 0.2 * JP - 0.2 * JQ - 0.2 * JR + 0.2 * JS + 0.2 * KL
 + 0.2 * KM - 0.2 * KN - 0.2 * KO - 0.2 * KP + 0.2 * KQ + 0.2 * KR - 0.2 * KS
 - 0.2 * LM - 0.2 * LN + 0.2 * LO + 0.2 * LP - 0.2 * LQ - 0.6 * LR - 0.2 * LS
 - 0.2 * MN - 0.2 * MO - 0.2 * MP + 0.2 * MQ - 0.2 * MR + 0.2 * MS - 0.6 * NO
 - 0.2 * NP + 0.2 * NQ - 0.2 * NR - 0.2 * NS + 0.2 * OP + 0.2 * OQ + 0.2 * OR
 - 0.2 * OS - 0.2 * PQ - 0.2 * PR - 0.2 * PS + 0.2 * QR - 0.6 * QS + 0.2 * RS
[D] = D + 0.2 * AB - 0.2 * AC - 0.2 * AE + 0.2 * AF - 0.2 * AG + 0.2 * AH
 + 0.2 * AI - 0.2 * AJ + 0.2 * AK - 0.2 * AL - 0.2 * AM + 0.2 * AN - 0.2 * AO
 - 0.2 * AP - 0.2 * AQ - 0.6 * AR + 0.2 * AS - 0.2 * BC - 0.2 * BE - 0.2 * BF
 - 0.6 * BG - 0.2 * BH + 0.2 * BI - 0.2 * BJ - 0.2 * BK - 0.2 * BL + 0.2 * BM
 + 0.2 * BN + 0.2 * BO - 0.2 * BP + 0.2 * BQ + 0.2 * BR - 0.2 * BS - 0.2 * CE
 + 0.2 * CF + 0.2 * CG + 0.2 * CH - 0.2 * CI + 0.2 * CJ - 0.6 * CK - 0.2 * CL
 - 0.2 * CM + 0.2 * CN + 0.2 * CO - 0.2 * CP + 0.2 * CQ - 0.2 * CR - 0.2 * CS
 - 0.2 * EF + 0.2 * EG + 0.2 * EH + 0.2 * EI - 0.2 * EJ + 0.2 * EK - 0.6 * EL
 - 0.2 * EM - 0.2 * EN + 0.2 * EO + 0.2 * EP - 0.2 * EQ + 0.2 * ER - 0.2 * ES
 - 0.2 * FG - 0.2 * FH - 0.6 * FI - 0.2 * FJ + 0.2 * FK - 0.2 * FL - 0.2 * FM
 - 0.2 * FN + 0.2 * FO + 0.2 * FP + 0.2 * FQ - 0.2 * FR + 0.2 * FS - 0.2 * GH
 + 0.2 * GI - 0.2 * GJ + 0.2 * GK + 0.2 * GL - 0.2 * GM + 0.2 * GN - 0.2 * GO
 - 0.2 * GP + 0.2 * GQ - 0.2 * GR - 0.2 * GS - 0.2 * HI + 0.2 * HJ - 0.2 * HK
 - 0.2 * HL + 0.2 * HM + 0.2 * HN - 0.2 * HO + 0.2 * HP - 0.6 * HQ - 0.2 * HR
 - 0.2 * HS + 0.2 * IJ - 0.2 * IK - 0.2 * IL - 0.2 * IM - 0.2 * IN - 0.2 * IO
 - 0.2 * IP + 0.2 * IQ - 0.2 * IR + 0.2 * IS - 0.2 * JK + 0.2 * JL + 0.2 * JM
 - 0.6 * JN + 0.2 * JO - 0.2 * JP - 0.2 * JQ - 0.2 * JR + 0.2 * JS + 0.2 * KL
 + 0.2 * KM - 0.2 * KN + 0.2 * KO - 0.2 * KP - 0.2 * KQ - 0.2 * KR - 0.2 * KS
 + 0.2 * LM + 0.2 * LN - 0.2 * LO - 0.2 * LP - 0.2 * LQ + 0.2 * LR + 0.2 * LS
 - 0.2 * MN - 0.2 * MO + 0.2 * MP + 0.2 * MQ - 0.2 * MR - 0.6 * MS - 0.2 * NO
 - 0.2 * NP - 0.2 * NQ + 0.2 * NR - 0.2 * NS - 0.6 * OP - 0.2 * OQ + 0.2 * OR
 - 0.2 * OS + 0.2 * PQ + 0.2 * PR + 0.2 * PS - 0.2 * QR - 0.2 * QS + 0.2 * RS
[E] = E + 0.2 * AB - 0.2 * AC - 0.2 * AD - 0.2 * AF + 0.2 * AG - 0.2 * AH
 - 0.2 * AI + 0.2 * AJ + 0.2 * AK - 0.2 * AL + 0.2 * AM - 0.6 * AN - 0.2 * AO
 - 0.2 * AP + 0.2 * AQ - 0.2 * AR + 0.2 * AS + 0.2 * BC - 0.2 * BD - 0.2 * BF
 + 0.2 * BG - 0.2 * BH + 0.2 * BI + 0.2 * BJ - 0.2 * BK + 0.2 * BL - 0.2 * BM
 - 0.2 * BN + 0.2 * BO - 0.2 * BP - 0.2 * BQ - 0.2 * BR - 0.6 * BS - 0.2 * CD
 - 0.2 * CF - 0.2 * CG - 0.6 * CH - 0.2 * CI + 0.2 * CJ - 0.2 * CK - 0.2 * CL
 - 0.2 * CM + 0.2 * CN + 0.2 * CO + 0.2 * CP - 0.2 * CQ + 0.2 * CR + 0.2 * CS
 - 0.2 * DF + 0.2 * DG + 0.2 * DH + 0.2 * DI - 0.2 * DJ + 0.2 * DK - 0.6 * DL
 - 0.2 * DM - 0.2 * DN + 0.2 * DO + 0.2 * DP - 0.2 * DQ + 0.2 * DR - 0.2 * DS
 - 0.2 * FG + 0.2 * FH + 0.2 * FI + 0.2 * FJ - 0.2 * FK + 0.2 * FL - 0.6 * FM
 - 0.2 * FN - 0.2 * FO + 0.2 * FP + 0.2 * FQ - 0.2 * FR + 0.2 * FS - 0.2 * GH
 - 0.2 * GI - 0.6 * GJ - 0.2 * GK + 0.2 * GL - 0.2 * GM - 0.2 * GN - 0.2 * GO
 + 0.2 * GP + 0.2 * GQ + 0.2 * GR - 0.2 * GS - 0.2 * HI + 0.2 * HJ - 0.2 * HK
 + 0.2 * HL + 0.2 * HM - 0.2 * HN + 0.2 * HO - 0.2 * HP - 0.2 * HQ + 0.2 * HR
 - 0.2 * HS - 0.2 * IJ + 0.2 * IK - 0.2 * IL - 0.2 * IM + 0.2 * IN + 0.2 * IO
 - 0.2 * IP + 0.2 * IQ - 0.6 * IR - 0.2 * IS + 0.2 * JK - 0.2 * JL - 0.2 * JM
 - 0.2 * JN - 0.2 * JO - 0.2 * JP - 0.2 * JQ + 0.2 * JR - 0.2 * JS - 0.2 * KL
 + 0.2 * KM + 0.2 * KN - 0.6 * KO + 0.2 * KP - 0.2 * KQ - 0.2 * KR - 0.2 * KS
 + 0.2 * LM + 0.2 * LN - 0.2 * LO + 0.2 * LP - 0.2 * LQ - 0.2 * LR - 0.2 * LS
 + 0.2 * MN + 0.2 * MO - 0.2 * MP - 0.2 * MQ - 0.2 * MR + 0.2 * MS - 0.2 * NO
 - 0.2 * NP + 0.2 * NQ + 0.2 * NR - 0.2 * NS - 0.2 * OP - 0.2 * OQ - 0.2 * OR
 + 0.2 * OS - 0.6 * PQ - 0.2 * PR + 0.2 * PS + 0.2 * QR + 0.2 * QS - 0.2 * RS
[F] = F + 0.2 * AB - 0.6 * AC + 0.2 * AD - 0.2 * AE + 0.2 * AG - 0.2 * AH
 - 0.2 * AI - 0.2 * AJ - 0.2 * AK - 0.2 * AL - 0.2 * AM + 0.2 * AN - 0.2 * AO
 + 0.2 * AP + 0.2 * AQ + 0.2 * AR - 0.2 * AS + 0.2 * BC - 0.2 * BD - 0.2 * BE
 - 0.2 * BG + 0.2 * BH - 0.2 * BI - 0.2 * BJ + 0.2 * BK + 0.2 * BL - 0.2 * BM
 + 0.2 * BN - 0.6 * BO - 0.2 * BP - 0.2 * BQ + 0.2 * BR - 0.2 * BS + 0.2 * CD
 - 0.2 * CE - 0.2 * CG + 0.2 * CH - 0.2 * CI + 0.2 * CJ + 0.2 * CK - 0.2 * CL
 + 0.2 * CM - 0.2 * CN - 0.2 * CO + 0.2 * CP - 0.2 * CQ - 0.2 * CR - 0.2 * CS
 - 0.2 * DE - 0.2 * DG - 0.2 * DH - 0.6 * DI - 0.2 * DJ + 0.2 * DK - 0.2 * DL
 - 0.2 * DM - 0.2 * DN + 0.2 * DO + 0.2 * DP + 0.2 * DQ - 0.2 * DR + 0.2 * DS
 - 0.2 * EG + 0.2 * EH + 0.2 * EI + 0.2 * EJ - 0.2 * EK + 0.2 * EL - 0.6 * EM
 - 0.2 * EN - 0.2 * EO + 0.2 * EP + 0.2 * EQ - 0.2 * ER + 0.2 * ES - 0.2 * GH
 + 0.2 * GI + 0.2 * GJ + 0.2 * GK - 0.2 * GL + 0.2 * GM - 0.6 * GN - 0.2 * GO
```

```
 - 0.2 * GP + 0.2 * GQ + 0.2 * GR - 0.2 * GS - 0.2 * HI - 0.2 * HJ - 0.6 * HK
 - 0.2 * HL + 0.2 * HM - 0.2 * HN - 0.2 * HO - 0.2 * HP + 0.2 * HQ + 0.2 * HR
 + 0.2 * HS - 0.2 * IJ + 0.2 * IK - 0.2 * IL + 0.2 * IM + 0.2 * IN - 0.2 * IO
 + 0.2 * IP - 0.2 * IQ - 0.2 * IR + 0.2 * IS - 0.2 * JK + 0.2 * JL - 0.2 * JM
 - 0.2 * JN + 0.2 * JO + 0.2 * JP - 0.2 * JQ + 0.2 * JR - 0.6 * JS + 0.2 * KL
 - 0.2 * KM - 0.2 * KN - 0.2 * KO - 0.2 * KP - 0.2 * KQ - 0.2 * KR + 0.2 * KS
 - 0.2 * LM + 0.2 * LN + 0.2 * LO - 0.6 * LP + 0.2 * LQ - 0.2 * LR - 0.2 * LS
 + 0.2 * MN + 0.2 * MO - 0.2 * MP + 0.2 * MQ - 0.2 * MR - 0.2 * MS + 0.2 * NO
 + 0.2 * NP - 0.2 * NQ - 0.2 * NR - 0.2 * NS - 0.2 * OP - 0.2 * OQ + 0.2 * OR
 + 0.2 * OS - 0.2 * PQ - 0.2 * PR - 0.2 * PS - 0.6 * QR - 0.2 * QS + 0.2 * RS
[G] = G - 0.2 * AB - 0.2 * AC - 0.2 * AD + 0.2 * AE + 0.2 * AF - 0.2 * AH
 + 0.2 * AI + 0.2 * AJ - 0.6 * AK + 0.2 * AL - 0.2 * AM - 0.2 * AN - 0.2 * AO
 + 0.2 * AP - 0.2 * AQ - 0.2 * AR + 0.2 * AS + 0.2 * BC - 0.6 * BD + 0.2 * BE
 - 0.2 * BF + 0.2 * BH - 0.2 * BI - 0.2 * BJ - 0.2 * BK - 0.2 * BL - 0.2 * BM
 - 0.2 * BN + 0.2 * BO - 0.2 * BP + 0.2 * BQ + 0.2 * BR + 0.2 * BS + 0.2 * CD
 - 0.2 * CE - 0.2 * CF - 0.2 * CH + 0.2 * CI - 0.2 * CJ - 0.2 * CK + 0.2 * CL
 + 0.2 * CM - 0.2 * CN + 0.2 * CO - 0.6 * CP - 0.2 * CQ - 0.2 * CR + 0.2 * CS
 + 0.2 * DE - 0.2 * DF - 0.2 * DH + 0.2 * DI - 0.2 * DJ + 0.2 * DK + 0.2 * DL
 - 0.2 * DM + 0.2 * DN - 0.2 * DO - 0.2 * DP + 0.2 * DQ - 0.2 * DR - 0.2 * DS
 - 0.2 * EF - 0.2 * EH - 0.2 * EI - 0.6 * EJ - 0.2 * EK + 0.2 * EL - 0.2 * EM
 - 0.2 * EN - 0.2 * EO + 0.2 * EP + 0.2 * EQ + 0.2 * ER - 0.2 * ES - 0.2 * FH
 + 0.2 * FI + 0.2 * FJ + 0.2 * FK - 0.2 * FL + 0.2 * FM - 0.6 * FN - 0.2 * FO
 - 0.2 * FP + 0.2 * FQ + 0.2 * FR - 0.2 * FS - 0.2 * HI + 0.2 * HJ + 0.2 * HK
 + 0.2 * HL - 0.2 * HM + 0.2 * HN - 0.6 * HO - 0.2 * HP - 0.2 * HQ + 0.2 * HR
 + 0.2 * HS - 0.2 * IJ - 0.2 * IK - 0.6 * IL - 0.2 * IM + 0.2 * IN - 0.2 * IO
 - 0.2 * IP - 0.2 * IQ + 0.2 * IR + 0.2 * IS - 0.2 * JK + 0.2 * JL - 0.2 * JM
 + 0.2 * JN + 0.2 * JO - 0.2 * JP + 0.2 * JQ - 0.2 * JR - 0.2 * JS - 0.2 * KL
 + 0.2 * KM - 0.2 * KN - 0.2 * KO + 0.2 * KP + 0.2 * KQ - 0.2 * KR + 0.2 * KS
 + 0.2 * LM - 0.2 * LN - 0.2 * LO - 0.2 * LP - 0.2 * LQ - 0.2 * LR - 0.2 * LS
 - 0.2 * MN + 0.2 * MO + 0.2 * MP - 0.6 * MQ + 0.2 * MR - 0.2 * MS + 0.2 * NO
 + 0.2 * NP - 0.2 * NQ + 0.2 * NR - 0.2 * NS + 0.2 * OP + 0.2 * OQ - 0.2 * OR
 - 0.2 * OS - 0.2 * PQ - 0.2 * PR + 0.2 * PS - 0.2 * QR - 0.2 * QS - 0.6 * RS
[H] = H + 0.2 * AB + 0.2 * AC + 0.2 * AD - 0.2 * AE - 0.2 * AF - 0.2 * AG
 + 0.2 * AI + 0.2 * AJ - 0.2 * AK + 0.2 * AL - 0.2 * AM - 0.2 * AN - 0.2 * AO
 - 0.2 * AP + 0.2 * AQ - 0.2 * AR - 0.6 * AS - 0.2 * BC - 0.2 * BD - 0.2 * BE
 + 0.2 * BF + 0.2 * BG - 0.2 * BI + 0.2 * BJ + 0.2 * BK - 0.6 * BL + 0.2 * BM
 - 0.2 * BN - 0.2 * BO - 0.2 * BP + 0.2 * BQ - 0.2 * BR - 0.2 * BS + 0.2 * CD
 - 0.6 * CE + 0.2 * CF - 0.2 * CG + 0.2 * CI - 0.2 * CJ - 0.2 * CK - 0.2 * CL
 - 0.2 * CM - 0.2 * CN - 0.2 * CO + 0.2 * CP - 0.2 * CQ + 0.2 * CR + 0.2 * CS
 + 0.2 * DE - 0.2 * DF - 0.2 * DG - 0.2 * DI + 0.2 * DJ - 0.2 * DK - 0.2 * DL
 + 0.2 * DM + 0.2 * DN - 0.2 * DO + 0.2 * DP - 0.6 * DQ - 0.2 * DR - 0.2 * DS
 + 0.2 * EF - 0.2 * EG - 0.2 * EI + 0.2 * EJ - 0.2 * EK + 0.2 * EL + 0.2 * EM
 - 0.2 * EN + 0.2 * EO - 0.2 * EP - 0.2 * EQ + 0.2 * ER - 0.2 * ES - 0.2 * FG
 - 0.2 * FI - 0.2 * FJ - 0.6 * FK - 0.2 * FL + 0.2 * FM - 0.2 * FN - 0.2 * FO
 - 0.2 * FP + 0.2 * FQ + 0.2 * FR + 0.2 * FS - 0.2 * GI + 0.2 * GJ + 0.2 * GK
 + 0.2 * GL - 0.2 * GM + 0.2 * GN - 0.6 * GO - 0.2 * GP - 0.2 * GQ + 0.2 * GR
 + 0.2 * GS - 0.2 * IJ + 0.2 * IK + 0.2 * IL + 0.2 * IM - 0.2 * IN + 0.2 * IO
 - 0.6 * IP - 0.2 * IQ - 0.2 * IR + 0.2 * IS - 0.2 * JK - 0.2 * JL - 0.6 * JM
 - 0.2 * JN + 0.2 * JO - 0.2 * JP - 0.2 * JQ - 0.2 * JR + 0.2 * JS - 0.2 * KL
 + 0.2 * KM - 0.2 * KN + 0.2 * KO + 0.2 * KP - 0.2 * KQ + 0.2 * KR - 0.2 * KS
 - 0.2 * LM + 0.2 * LN - 0.2 * LO - 0.2 * LP + 0.2 * LQ + 0.2 * LR - 0.2 * LS
 + 0.2 * MN - 0.2 * MO - 0.2 * MP - 0.2 * MQ - 0.2 * MR - 0.2 * MS - 0.2 * NO
 + 0.2 * NP + 0.2 * NQ - 0.6 * NR + 0.2 * NS + 0.2 * OP + 0.2 * OQ - 0.2 * OR
 + 0.2 * OS + 0.2 * PQ + 0.2 * PR - 0.2 * PS - 0.2 * QR - 0.2 * QS - 0.2 * RS
[I] = I - 0.6 * AB - 0.2 * AC + 0.2 * AD - 0.2 * AE - 0.2 * AF + 0.2 * AG
 + 0.2 * AH + 0.2 * AJ + 0.2 * AK - 0.2 * AL - 0.2 * AM - 0.2 * AN + 0.2 * AO
 + 0.2 * AP - 0.2 * AQ - 0.2 * AR - 0.2 * AS + 0.2 * BC + 0.2 * BD + 0.2 * BE
 - 0.2 * BF - 0.2 * BG - 0.2 * BH + 0.2 * BJ + 0.2 * BK - 0.2 * BL + 0.2 * BM
 - 0.2 * BN - 0.2 * BO - 0.2 * BP - 0.2 * BQ + 0.2 * BR - 0.2 * BS - 0.2 * CD
 - 0.2 * CE - 0.2 * CF + 0.2 * CG + 0.2 * CH - 0.2 * CJ + 0.2 * CK + 0.2 * CL
 - 0.6 * CM + 0.2 * CN - 0.2 * CO - 0.2 * CP - 0.2 * CQ + 0.2 * CR - 0.2 * CS
 + 0.2 * DE - 0.6 * DF + 0.2 * DG - 0.2 * DH + 0.2 * DJ - 0.2 * DK - 0.2 * DL
 - 0.2 * DM - 0.2 * DN - 0.2 * DO - 0.2 * DP + 0.2 * DQ - 0.2 * DR + 0.2 * DS
 + 0.2 * EF - 0.2 * EG - 0.2 * EH - 0.2 * EJ + 0.2 * EK - 0.2 * EL - 0.2 * EM
 + 0.2 * EN + 0.2 * EO - 0.2 * EP + 0.2 * EQ - 0.6 * ER - 0.2 * ES + 0.2 * FG
 - 0.2 * FH - 0.2 * FJ + 0.2 * FK - 0.2 * FL + 0.2 * FM + 0.2 * FN - 0.2 * FO
 + 0.2 * FP - 0.2 * FQ - 0.2 * FR + 0.2 * FS - 0.2 * GH - 0.2 * GJ - 0.2 * GK
 - 0.6 * GL - 0.2 * GM + 0.2 * GN - 0.2 * GO - 0.2 * GP - 0.2 * GQ + 0.2 * GR
 + 0.2 * GS - 0.2 * HJ + 0.2 * HK + 0.2 * HL + 0.2 * HM - 0.2 * HN + 0.2 * HO
 - 0.6 * HP - 0.2 * HQ - 0.2 * HR + 0.2 * HS - 0.2 * JK + 0.2 * JL + 0.2 * JM
```

```
  + 0.2 * JN - 0.2 * JO + 0.2 * JP - 0.6 * JQ - 0.2 * JR - 0.2 * JS - 0.2 * KL
  - 0.2 * KM - 0.6 * KN - 0.2 * KO + 0.2 * KP - 0.2 * KQ - 0.2 * KR - 0.2 * KS
  - 0.2 * LM + 0.2 * LN - 0.2 * LO + 0.2 * LP + 0.2 * LQ - 0.2 * LR + 0.2 * LS
  - 0.2 * MN + 0.2 * MO - 0.2 * MP - 0.2 * MQ + 0.2 * MR + 0.2 * MS + 0.2 * NO
  - 0.2 * NP - 0.2 * NQ - 0.2 * NR - 0.2 * NS - 0.2 * OP + 0.2 * OQ + 0.2 * OR
  - 0.6 * OS + 0.2 * PQ + 0.2 * PR - 0.2 * PS + 0.2 * QR + 0.2 * QS - 0.2 * RS
[J] = J - 0.2 * AB - 0.2 * AC - 0.2 * AD + 0.2 * AE - 0.2 * AF + 0.2 * AG
  + 0.2 * AH + 0.2 * AI - 0.2 * AK - 0.2 * AL + 0.2 * AM + 0.2 * AN - 0.2 * AO
  - 0.6 * AP - 0.2 * AQ + 0.2 * AR - 0.2 * AS - 0.6 * BC - 0.2 * BD + 0.2 * BE
  - 0.2 * BF - 0.2 * BG + 0.2 * BH + 0.2 * BI + 0.2 * BK + 0.2 * BL - 0.2 * BM
  - 0.2 * BN - 0.2 * BO + 0.2 * BP + 0.2 * BQ - 0.2 * BR - 0.2 * BS + 0.2 * CD
  + 0.2 * CE + 0.2 * CF - 0.2 * CG - 0.2 * CH - 0.2 * CI + 0.2 * CK + 0.2 * CL
  - 0.2 * CM + 0.2 * CN - 0.2 * CO - 0.2 * CP - 0.2 * CQ - 0.2 * CR + 0.2 * CS
  - 0.2 * DE - 0.2 * DF - 0.2 * DG + 0.2 * DH + 0.2 * DI - 0.2 * DK + 0.2 * DL
  + 0.2 * DM - 0.6 * DN + 0.2 * DO - 0.2 * DP - 0.2 * DQ - 0.2 * DR + 0.2 * DS
  + 0.2 * EF - 0.6 * EG + 0.2 * EH - 0.2 * EI + 0.2 * EK - 0.2 * EL - 0.2 * EM
  - 0.2 * EN - 0.2 * EO - 0.2 * EP - 0.2 * EQ + 0.2 * ER - 0.2 * ES + 0.2 * FG
  - 0.2 * FH - 0.2 * FI - 0.2 * FK + 0.2 * FL - 0.2 * FM - 0.2 * FN + 0.2 * FO
  + 0.2 * FP - 0.2 * FQ + 0.2 * FR - 0.6 * FS + 0.2 * GH - 0.2 * GI - 0.2 * GK
  + 0.2 * GL - 0.2 * GM + 0.2 * GN + 0.2 * GO - 0.2 * GP + 0.2 * GQ - 0.2 * GR
  - 0.2 * GS - 0.2 * HI - 0.2 * HK - 0.2 * HL - 0.6 * HM - 0.2 * HN + 0.2 * HO
  - 0.2 * HP - 0.2 * HQ - 0.2 * HR + 0.2 * HS - 0.2 * IK + 0.2 * IL + 0.2 * IM
  + 0.2 * IN - 0.2 * IO + 0.2 * IP - 0.6 * IQ - 0.2 * IR - 0.2 * IS - 0.2 * KL
  + 0.2 * KM + 0.2 * KN + 0.2 * KO - 0.2 * KP + 0.2 * KQ - 0.6 * KR - 0.2 * KS
  - 0.2 * LM - 0.2 * LN - 0.6 * LO - 0.2 * LP + 0.2 * LQ - 0.2 * LR - 0.2 * LS
  - 0.2 * MN + 0.2 * MO - 0.2 * MP + 0.2 * MQ + 0.2 * MR - 0.2 * MS - 0.2 * NO
  + 0.2 * NP - 0.2 * NQ - 0.2 * NR + 0.2 * NS + 0.2 * OP - 0.2 * OQ - 0.2 * OR
  - 0.2 * OS - 0.2 * PQ + 0.2 * PR + 0.2 * PS + 0.2 * QR + 0.2 * QS + 0.2 * RS
[K] = K - 0.2 * AB - 0.2 * AC + 0.2 * AD + 0.2 * AE - 0.2 * AF - 0.6 * AG
  - 0.2 * AH + 0.2 * AI - 0.2 * AJ - 0.2 * AL - 0.2 * AM - 0.2 * AN + 0.2 * AO
  - 0.2 * AP + 0.2 * AQ + 0.2 * AR + 0.2 * AS - 0.2 * BC - 0.2 * BD - 0.2 * BE
  + 0.2 * BF - 0.2 * BG + 0.2 * BH + 0.2 * BI + 0.2 * BJ - 0.2 * BL - 0.2 * BM
  + 0.2 * BN + 0.2 * BO - 0.2 * BP - 0.6 * BQ - 0.2 * BR + 0.2 * BS - 0.6 * CD
  - 0.2 * CE + 0.2 * CF - 0.2 * CG - 0.2 * CH + 0.2 * CI + 0.2 * CJ + 0.2 * CL
  + 0.2 * CM - 0.2 * CN - 0.2 * CO - 0.2 * CP + 0.2 * CQ + 0.2 * CR - 0.2 * CS
  + 0.2 * DE + 0.2 * DF + 0.2 * DG - 0.2 * DH - 0.2 * DI - 0.2 * DJ + 0.2 * DL
  + 0.2 * DM - 0.2 * DN + 0.2 * DO - 0.2 * DP - 0.2 * DQ - 0.2 * DR - 0.2 * DS
  - 0.2 * EF - 0.2 * EG - 0.2 * EH + 0.2 * EI + 0.2 * EJ - 0.2 * EL + 0.2 * EM
  + 0.2 * EN - 0.6 * EO + 0.2 * EP - 0.2 * EQ - 0.2 * ER - 0.2 * ES + 0.2 * FG
  - 0.6 * FH + 0.2 * FI - 0.2 * FJ + 0.2 * FL - 0.2 * FM - 0.2 * FN - 0.2 * FO
  - 0.2 * FP - 0.2 * FQ - 0.2 * FR + 0.2 * FS + 0.2 * GH - 0.2 * GI - 0.2 * GJ
  - 0.2 * GL + 0.2 * GM - 0.2 * GN - 0.2 * GO + 0.2 * GP + 0.2 * GQ - 0.2 * GR
  + 0.2 * GS + 0.2 * HI - 0.2 * HJ - 0.2 * HL + 0.2 * HM - 0.2 * HN + 0.2 * HO
  + 0.2 * HP - 0.2 * HQ + 0.2 * HR - 0.2 * HS - 0.2 * IJ - 0.2 * IL - 0.2 * IM
  - 0.6 * IN - 0.2 * IO + 0.2 * IP - 0.2 * IQ - 0.2 * IR - 0.2 * IS - 0.2 * JL
  + 0.2 * JM + 0.2 * JN + 0.2 * JO - 0.2 * JP + 0.2 * JQ - 0.6 * JR - 0.2 * JS
  - 0.2 * LM + 0.2 * LN + 0.2 * LO + 0.2 * LP - 0.2 * LQ + 0.2 * LR - 0.6 * LS
  - 0.2 * MN - 0.2 * MO - 0.6 * MP - 0.2 * MQ + 0.2 * MR - 0.2 * MS - 0.2 * NO
  + 0.2 * NP - 0.2 * NQ + 0.2 * NR + 0.2 * NS - 0.2 * OP + 0.2 * OQ - 0.2 * OR
  - 0.2 * OS + 0.2 * PQ - 0.2 * PR - 0.2 * PS - 0.2 * QR + 0.2 * QS + 0.2 * RS
[L] = L + 0.2 * AB + 0.2 * AC - 0.2 * AD - 0.2 * AE - 0.2 * AF + 0.2 * AG
  + 0.2 * AH - 0.2 * AI - 0.2 * AJ - 0.2 * AK - 0.6 * AM - 0.2 * AN + 0.2 * AO
  - 0.2 * AP - 0.2 * AQ + 0.2 * AR + 0.2 * AS - 0.2 * BC - 0.2 * BD + 0.2 * BE
  + 0.2 * BF - 0.2 * BG - 0.6 * BH - 0.2 * BI + 0.2 * BJ - 0.2 * BK - 0.2 * BM
  - 0.2 * BN - 0.2 * BO + 0.2 * BP - 0.2 * BQ + 0.2 * BR + 0.2 * BS - 0.2 * CD
  - 0.2 * CE - 0.2 * CF + 0.2 * CG - 0.2 * CH + 0.2 * CI + 0.2 * CJ + 0.2 * CK
  - 0.2 * CM - 0.2 * CN + 0.2 * CO + 0.2 * CP - 0.2 * CQ - 0.6 * CR - 0.2 * CS
  - 0.6 * DE - 0.2 * DF + 0.2 * DG - 0.2 * DH - 0.2 * DI + 0.2 * DJ + 0.2 * DK
  + 0.2 * DM + 0.2 * DN - 0.2 * DO - 0.2 * DP - 0.2 * DQ + 0.2 * DR + 0.2 * DS
  + 0.2 * EF + 0.2 * EG + 0.2 * EH - 0.2 * EI - 0.2 * EJ - 0.2 * EK + 0.2 * EM
  + 0.2 * EN - 0.2 * EO + 0.2 * EP - 0.2 * EQ - 0.2 * ER - 0.2 * ES - 0.2 * FG
  - 0.2 * FH - 0.2 * FI + 0.2 * FJ + 0.2 * FK - 0.2 * FM + 0.2 * FN + 0.2 * FO
  - 0.6 * FP + 0.2 * FQ - 0.2 * FR - 0.2 * FS + 0.2 * GH - 0.6 * GI + 0.2 * GJ
  - 0.2 * GK + 0.2 * GM - 0.2 * GN - 0.2 * GO - 0.2 * GP - 0.2 * GQ - 0.2 * GR
  - 0.2 * GS + 0.2 * HI - 0.2 * HJ - 0.2 * HK - 0.2 * HM + 0.2 * HN - 0.2 * HO
  - 0.2 * HP + 0.2 * HQ + 0.2 * HR - 0.2 * HS + 0.2 * IJ - 0.2 * IK - 0.2 * IM
  + 0.2 * IN - 0.2 * IO + 0.2 * IP + 0.2 * IQ - 0.2 * IR + 0.2 * IS - 0.2 * JK
  - 0.2 * JM - 0.2 * JN - 0.6 * JO - 0.2 * JP + 0.2 * JQ - 0.2 * JR - 0.2 * JS
  - 0.2 * KM + 0.2 * KN + 0.2 * KO + 0.2 * KP - 0.2 * KQ + 0.2 * KR - 0.6 * KS
  - 0.2 * MN + 0.2 * MO + 0.2 * MP + 0.2 * MQ - 0.2 * MR + 0.2 * MS - 0.2 * NO
```

- 0.2 * NP - 0.6 * NQ - 0.2 * NR + 0.2 * NS - 0.2 * OP + 0.2 * OQ - 0.2 * OR
+ 0.2 * OS - 0.2 * PQ + 0.2 * PR - 0.2 * PS + 0.2 * QR - 0.2 * QS - 0.2 * RS
[M] = M + 0.2 * AB + 0.2 * AC - 0.2 * AD + 0.2 * AE - 0.2 * AF - 0.2 * AG
- 0.2 * AH - 0.2 * AI + 0.2 * AJ - 0.2 * AK - 0.6 * AL + 0.2 * AN + 0.2 * AO
+ 0.2 * AP - 0.2 * AQ - 0.2 * AR - 0.2 * AS + 0.2 * BC + 0.2 * BD - 0.2 * BE
- 0.2 * BF - 0.2 * BG + 0.2 * BH + 0.2 * BI - 0.2 * BJ - 0.2 * BK - 0.2 * BL
- 0.6 * BN - 0.2 * BO + 0.2 * BP - 0.2 * BQ - 0.2 * BR + 0.2 * BS - 0.2 * CD
- 0.2 * CE + 0.2 * CF + 0.2 * CG - 0.2 * CH - 0.6 * CI - 0.2 * CJ + 0.2 * CK
- 0.2 * CL - 0.2 * CN - 0.2 * CO - 0.2 * CP + 0.2 * CQ - 0.2 * CR + 0.2 * CS
- 0.2 * DE - 0.2 * DF - 0.2 * DG + 0.2 * DH - 0.2 * DI + 0.2 * DJ + 0.2 * DK
+ 0.2 * DL - 0.2 * DN - 0.2 * DO + 0.2 * DP + 0.2 * DQ - 0.2 * DR - 0.6 * DS
- 0.6 * EF - 0.2 * EG + 0.2 * EH - 0.2 * EI - 0.2 * EJ + 0.2 * EK + 0.2 * EL
+ 0.2 * EN + 0.2 * EO - 0.2 * EP - 0.2 * EQ - 0.2 * ER + 0.2 * ES + 0.2 * FG
+ 0.2 * FH + 0.2 * FI - 0.2 * FJ - 0.2 * FK - 0.2 * FL + 0.2 * FN + 0.2 * FO
- 0.2 * FP + 0.2 * FQ - 0.2 * FR - 0.2 * FS - 0.2 * GH - 0.2 * GI - 0.2 * GJ
+ 0.2 * GK + 0.2 * GL - 0.2 * GN + 0.2 * GO + 0.2 * GP - 0.6 * GQ + 0.2 * GR
- 0.2 * GS + 0.2 * HI - 0.6 * HJ + 0.2 * HK - 0.2 * HL + 0.2 * HN - 0.2 * HO
- 0.2 * HP - 0.2 * HQ - 0.2 * HR - 0.2 * HS + 0.2 * IJ - 0.2 * IK - 0.2 * IL
- 0.2 * IN + 0.2 * IO - 0.2 * IP - 0.2 * IQ + 0.2 * IR + 0.2 * IS + 0.2 * JK
- 0.2 * JL - 0.2 * JN + 0.2 * JO - 0.2 * JP + 0.2 * JQ + 0.2 * JR - 0.2 * JS
- 0.2 * KL - 0.2 * KN - 0.2 * KO - 0.6 * KP - 0.2 * KQ + 0.2 * KR - 0.2 * KS
- 0.2 * LN + 0.2 * LO + 0.2 * LP + 0.2 * LQ - 0.2 * LR + 0.2 * LS - 0.2 * NO
+ 0.2 * NP + 0.2 * NQ + 0.2 * NR - 0.2 * NS - 0.2 * OP - 0.2 * OQ - 0.6 * OR
- 0.2 * OS - 0.2 * PQ + 0.2 * PR - 0.2 * PS - 0.2 * QR + 0.2 * QS + 0.2 * RS
[N] = N - 0.2 * AB + 0.2 * AC + 0.2 * AD - 0.6 * AE + 0.2 * AF - 0.2 * AG
- 0.2 * AH - 0.2 * AI + 0.2 * AJ - 0.2 * AK - 0.2 * AL + 0.2 * AM - 0.2 * AO
- 0.2 * AP - 0.2 * AQ + 0.2 * AR + 0.2 * AS + 0.2 * BC + 0.2 * BD - 0.2 * BE
+ 0.2 * BF - 0.2 * BG - 0.2 * BH - 0.2 * BI - 0.2 * BJ + 0.2 * BK - 0.2 * BL
- 0.6 * BM + 0.2 * BO + 0.2 * BP + 0.2 * BQ - 0.2 * BR - 0.2 * BS + 0.2 * CD
+ 0.2 * CE - 0.2 * CF - 0.2 * CG - 0.2 * CH + 0.2 * CI + 0.2 * CJ - 0.2 * CK
- 0.2 * CL - 0.2 * CM - 0.6 * CO - 0.2 * CP + 0.2 * CQ - 0.2 * CR - 0.2 * CS
- 0.2 * DE - 0.2 * DF + 0.2 * DG + 0.2 * DH - 0.2 * DI - 0.6 * DJ - 0.2 * DK
+ 0.2 * DL - 0.2 * DM - 0.2 * DO - 0.2 * DP - 0.2 * DQ + 0.2 * DR - 0.2 * DS
- 0.2 * EF - 0.2 * EG - 0.2 * EH + 0.2 * EI - 0.2 * EJ + 0.2 * EK + 0.2 * EL
+ 0.2 * EM - 0.2 * EO - 0.2 * EP + 0.2 * EQ + 0.2 * ER - 0.2 * ES - 0.6 * FG
- 0.2 * FH + 0.2 * FI - 0.2 * FJ - 0.2 * FK + 0.2 * FL + 0.2 * FM + 0.2 * FO
+ 0.2 * FP - 0.2 * FQ - 0.2 * FR - 0.2 * FS + 0.2 * GH + 0.2 * GI + 0.2 * GJ
- 0.2 * GK - 0.2 * GL - 0.2 * GM + 0.2 * GO + 0.2 * GP - 0.2 * GQ + 0.2 * GR
- 0.2 * GS - 0.2 * HI - 0.2 * HJ - 0.2 * HK + 0.2 * HL + 0.2 * HM - 0.2 * HO
+ 0.2 * HP + 0.2 * HQ - 0.6 * HR + 0.2 * HS + 0.2 * IJ - 0.6 * IK + 0.2 * IL
- 0.2 * IM + 0.2 * IO - 0.2 * IP - 0.2 * IQ - 0.2 * IR - 0.2 * IS + 0.2 * JK
- 0.2 * JL - 0.2 * JM - 0.2 * JO + 0.2 * JP - 0.2 * JQ - 0.2 * JR + 0.2 * JS
+ 0.2 * KL - 0.2 * KM - 0.2 * KO + 0.2 * KP - 0.2 * KQ + 0.2 * KR + 0.2 * KS
- 0.2 * LM - 0.2 * LO - 0.2 * LP - 0.6 * LQ - 0.2 * LR + 0.2 * LS - 0.2 * MO
+ 0.2 * MP + 0.2 * MQ + 0.2 * MR - 0.2 * MS - 0.2 * OP + 0.2 * OQ + 0.2 * OR
+ 0.2 * OS - 0.2 * PQ - 0.2 * PR - 0.6 * PS - 0.2 * QR + 0.2 * QS - 0.2 * RS
[O] = O + 0.2 * AB - 0.2 * AC - 0.2 * AD - 0.2 * AE - 0.2 * AF - 0.2 * AG
- 0.2 * AH + 0.2 * AI - 0.2 * AJ + 0.2 * AK + 0.2 * AL + 0.2 * AM - 0.2 * AN
+ 0.2 * AP - 0.6 * AQ + 0.2 * AR - 0.2 * AS - 0.2 * BC + 0.2 * BD + 0.2 * BE
- 0.6 * BF + 0.2 * BG - 0.2 * BH - 0.2 * BI - 0.2 * BJ + 0.2 * BK - 0.2 * BL
- 0.2 * BM + 0.2 * BN - 0.2 * BP - 0.2 * BQ - 0.2 * BR + 0.2 * BS + 0.2 * CD
+ 0.2 * CE - 0.2 * CF + 0.2 * CG - 0.2 * CH - 0.2 * CI - 0.2 * CJ - 0.2 * CK
+ 0.2 * CL - 0.2 * CM - 0.6 * CN + 0.2 * CP + 0.2 * CQ + 0.2 * CR - 0.2 * CS
+ 0.2 * DE + 0.2 * DF - 0.2 * DG - 0.2 * DH - 0.2 * DI + 0.2 * DJ + 0.2 * DK
- 0.2 * DL - 0.2 * DM - 0.2 * DN - 0.6 * DP - 0.2 * DQ + 0.2 * DR - 0.2 * DS
- 0.2 * EF - 0.2 * EG + 0.2 * EH + 0.2 * EI - 0.2 * EJ - 0.6 * EK - 0.2 * EL
+ 0.2 * EM - 0.2 * EN - 0.2 * EP - 0.2 * EQ - 0.2 * ER + 0.2 * ES - 0.2 * FG
- 0.2 * FH - 0.2 * FI + 0.2 * FJ - 0.2 * FK + 0.2 * FL + 0.2 * FM + 0.2 * FN
- 0.2 * FP - 0.2 * FQ + 0.2 * FR + 0.2 * FS - 0.6 * GH - 0.2 * GI + 0.2 * GJ
- 0.2 * GK - 0.2 * GL + 0.2 * GM + 0.2 * GN + 0.2 * GP + 0.2 * GQ - 0.2 * GR
- 0.2 * GS + 0.2 * HI + 0.2 * HJ + 0.2 * HK - 0.2 * HL - 0.2 * HM - 0.2 * HN
+ 0.2 * HP + 0.2 * HQ - 0.2 * HR + 0.2 * HS - 0.2 * IJ - 0.2 * IK - 0.2 * IL
+ 0.2 * IM + 0.2 * IN - 0.2 * IP + 0.2 * IQ + 0.2 * IR - 0.6 * IS + 0.2 * JK
- 0.6 * JL + 0.2 * JM - 0.2 * JN + 0.2 * JP - 0.2 * JQ - 0.2 * JR - 0.2 * JS
+ 0.2 * KL - 0.2 * KM - 0.2 * KN - 0.2 * KP + 0.2 * KQ - 0.2 * KR - 0.2 * KS
+ 0.2 * LM - 0.2 * LN - 0.2 * LP + 0.2 * LQ - 0.2 * LR + 0.2 * LS - 0.2 * MN
- 0.2 * MP - 0.2 * MQ - 0.6 * MR - 0.2 * MS - 0.2 * NP + 0.2 * NQ + 0.2 * NR
+ 0.2 * NS - 0.2 * PQ + 0.2 * PR + 0.2 * PS - 0.2 * QR - 0.2 * QS - 0.2 * RS
[P] = P - 0.2 * AB + 0.2 * AC - 0.2 * AD - 0.2 * AE + 0.2 * AF + 0.2 * AG
- 0.2 * AH + 0.2 * AI - 0.6 * AJ - 0.2 * AK - 0.2 * AL + 0.2 * AM - 0.2 * AN

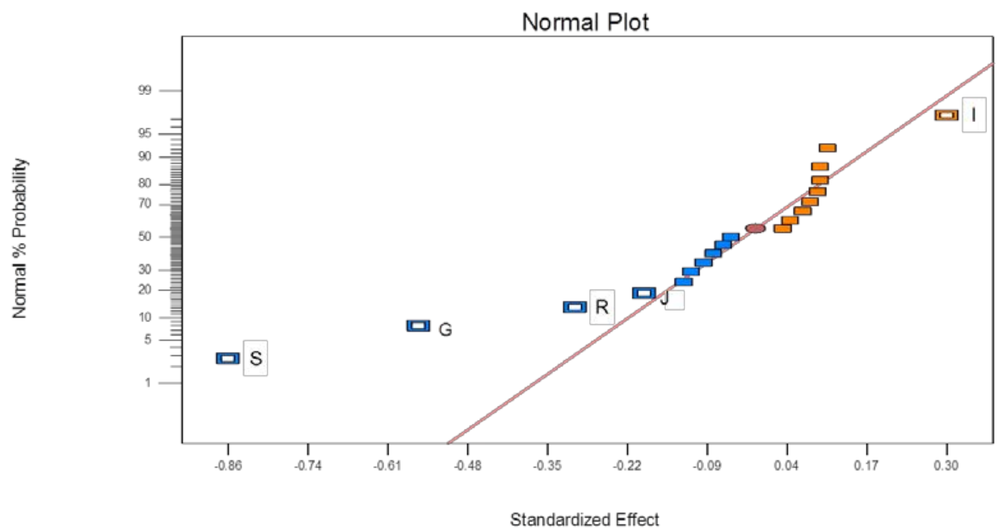
```
 + 0.2 * AO + 0.2 * AQ - 0.2 * AR - 0.2 * AS + 0.2 * BC - 0.2 * BD - 0.2 * BE
 - 0.2 * BF - 0.2 * BG - 0.2 * BH - 0.2 * BI + 0.2 * BJ - 0.2 * BK + 0.2 * BL
 + 0.2 * BM + 0.2 * BN - 0.2 * BO + 0.2 * BQ - 0.6 * BR + 0.2 * BS - 0.2 * CD
 + 0.2 * CE + 0.2 * CF - 0.6 * CG + 0.2 * CH - 0.2 * CI - 0.2 * CJ - 0.2 * CK
 + 0.2 * CL - 0.2 * CM - 0.2 * CN + 0.2 * CO - 0.2 * CQ - 0.2 * CR - 0.2 * CS
 + 0.2 * DE + 0.2 * DF - 0.2 * DG + 0.2 * DH - 0.2 * DI - 0.2 * DJ - 0.2 * DK
 - 0.2 * DL + 0.2 * DM - 0.2 * DN - 0.6 * DO + 0.2 * DQ + 0.2 * DR + 0.2 * DS
 + 0.2 * EF + 0.2 * EG - 0.2 * EH - 0.2 * EI - 0.2 * EJ + 0.2 * EK + 0.2 * EL
 - 0.2 * EM - 0.2 * EN - 0.2 * EO - 0.6 * EQ - 0.2 * ER + 0.2 * ES - 0.2 * FG
 - 0.2 * FH + 0.2 * FI + 0.2 * FJ - 0.2 * FK - 0.6 * FL - 0.2 * FM + 0.2 * FN
 - 0.2 * FO - 0.2 * FQ - 0.2 * FR - 0.2 * FS - 0.2 * GH - 0.2 * GI - 0.2 * GJ
 + 0.2 * GK - 0.2 * GL + 0.2 * GM + 0.2 * GN + 0.2 * GO - 0.2 * GQ - 0.2 * GR
 + 0.2 * GS - 0.6 * HI - 0.2 * HJ + 0.2 * HK - 0.2 * HL - 0.2 * HM + 0.2 * HN
 + 0.2 * HO + 0.2 * HQ + 0.2 * HR - 0.2 * HS + 0.2 * IJ + 0.2 * IK + 0.2 * IL
 - 0.2 * IM - 0.2 * IN - 0.2 * IO + 0.2 * IQ + 0.2 * IR - 0.2 * IS - 0.2 * JK
 - 0.2 * JL - 0.2 * JM + 0.2 * JN + 0.2 * JO - 0.2 * JQ + 0.2 * JR + 0.2 * JS
 + 0.2 * KL - 0.6 * KM + 0.2 * KN - 0.2 * KO + 0.2 * KQ - 0.2 * KR - 0.2 * KS
 + 0.2 * LM - 0.2 * LN - 0.2 * LO - 0.2 * LQ + 0.2 * LR - 0.2 * LS + 0.2 * MN
 - 0.2 * MO - 0.2 * MQ + 0.2 * MR - 0.2 * MS - 0.2 * NO - 0.2 * NQ - 0.2 * NR
 - 0.6 * NS - 0.2 * OQ + 0.2 * OR + 0.2 * OS - 0.2 * QR + 0.2 * QS - 0.2 * RS
[Q] = Q - 0.2 * AB + 0.2 * AC - 0.2 * AD + 0.2 * AE + 0.2 * AF - 0.2 * AG
 + 0.2 * AH - 0.2 * AI - 0.2 * AJ + 0.2 * AK - 0.2 * AL - 0.2 * AM - 0.2 * AN
 - 0.6 * AO + 0.2 * AP + 0.2 * AR - 0.2 * AS - 0.2 * BC + 0.2 * BD - 0.2 * BE
 - 0.2 * BF + 0.2 * BG + 0.2 * BH - 0.2 * BI + 0.2 * BJ - 0.6 * BK - 0.2 * BL
 - 0.2 * BM + 0.2 * BN - 0.2 * BO + 0.2 * BP + 0.2 * BR - 0.2 * BS + 0.2 * CD
 - 0.2 * CE - 0.2 * CF - 0.2 * CG - 0.2 * CH - 0.2 * CI - 0.2 * CJ + 0.2 * CK
 - 0.2 * CL + 0.2 * CM + 0.2 * CN + 0.2 * CO - 0.2 * CP + 0.2 * CR - 0.6 * CS
 - 0.2 * DE + 0.2 * DF + 0.2 * DG - 0.6 * DH + 0.2 * DI - 0.2 * DJ - 0.2 * DK
 - 0.2 * DL + 0.2 * DM - 0.2 * DN - 0.2 * DO + 0.2 * DP - 0.2 * DR - 0.2 * DS
 + 0.2 * EF + 0.2 * EG - 0.2 * EH + 0.2 * EI - 0.2 * EJ - 0.2 * EK - 0.2 * EL
 - 0.2 * EM + 0.2 * EN - 0.2 * EO - 0.6 * EP + 0.2 * ER + 0.2 * ES + 0.2 * FG
 + 0.2 * FH - 0.2 * FI - 0.2 * FJ - 0.2 * FK + 0.2 * FL + 0.2 * FM - 0.2 * FN
 - 0.2 * FO - 0.2 * FP - 0.6 * FR - 0.2 * FS - 0.2 * GH - 0.2 * GI + 0.2 * GJ
 + 0.2 * GK - 0.2 * GL - 0.6 * GM - 0.2 * GN + 0.2 * GO - 0.2 * GP - 0.2 * GR
 - 0.2 * GS - 0.2 * HI - 0.2 * HJ - 0.2 * HK + 0.2 * HL - 0.2 * HM + 0.2 * HN
 + 0.2 * HO + 0.2 * HP - 0.2 * HR - 0.2 * HS - 0.6 * IJ - 0.2 * IK + 0.2 * IL
 - 0.2 * IM - 0.2 * IN + 0.2 * IO + 0.2 * IP + 0.2 * IR + 0.2 * IS + 0.2 * JK
 + 0.2 * JL + 0.2 * JM - 0.2 * JN - 0.2 * JO - 0.2 * JP + 0.2 * JR + 0.2 * JS
 - 0.2 * KL - 0.2 * KM - 0.2 * KN + 0.2 * KO + 0.2 * KP - 0.2 * KR + 0.2 * KS
 + 0.2 * LM - 0.6 * LN + 0.2 * LO - 0.2 * LP + 0.2 * LR - 0.2 * LS + 0.2 * MN
 - 0.2 * MO - 0.2 * MP - 0.2 * MR + 0.2 * MS + 0.2 * NO - 0.2 * NP - 0.2 * NR
 + 0.2 * NS - 0.2 * OP - 0.2 * OR - 0.2 * OS - 0.2 * PR + 0.2 * PS - 0.2 * RS
[R] = R - 0.2 * AB - 0.2 * AC - 0.6 * AD - 0.2 * AE + 0.2 * AF - 0.2 * AG
 - 0.2 * AH - 0.2 * AI + 0.2 * AJ + 0.2 * AK + 0.2 * AL - 0.2 * AM + 0.2 * AN
 + 0.2 * AO - 0.2 * AP + 0.2 * AQ - 0.2 * AS - 0.2 * BC + 0.2 * BD - 0.2 * BE
 + 0.2 * BF + 0.2 * BG - 0.2 * BH + 0.2 * BI - 0.2 * BJ - 0.2 * BK + 0.2 * BL
 - 0.2 * BM - 0.2 * BN - 0.2 * BO - 0.6 * BP + 0.2 * BQ + 0.2 * BS - 0.2 * CD
 + 0.2 * CE - 0.2 * CF - 0.2 * CG + 0.2 * CH + 0.2 * CI - 0.2 * CJ + 0.2 * CK
 - 0.6 * CL - 0.2 * CM - 0.2 * CN + 0.2 * CO - 0.2 * CP + 0.2 * CQ + 0.2 * CS
 + 0.2 * DE - 0.2 * DF - 0.2 * DG - 0.2 * DH - 0.2 * DI - 0.2 * DJ - 0.2 * DK
 + 0.2 * DL - 0.2 * DM + 0.2 * DN + 0.2 * DO + 0.2 * DP - 0.2 * DQ + 0.2 * DS
 - 0.2 * EF + 0.2 * EG + 0.2 * EH - 0.6 * EI + 0.2 * EJ - 0.2 * EK - 0.2 * EL
 - 0.2 * EM + 0.2 * EN - 0.2 * EO - 0.2 * EP + 0.2 * EQ - 0.2 * ES + 0.2 * FG
 + 0.2 * FH - 0.2 * FI + 0.2 * FJ - 0.2 * FK - 0.2 * FL - 0.2 * FM - 0.2 * FN
 + 0.2 * FO - 0.2 * FP - 0.6 * FQ + 0.2 * FS + 0.2 * GH + 0.2 * GI - 0.2 * GJ
 - 0.2 * GK - 0.2 * GL + 0.2 * GM + 0.2 * GN - 0.2 * GO - 0.2 * GP - 0.2 * GQ
 - 0.6 * GS - 0.2 * HI - 0.2 * HJ + 0.2 * HK + 0.2 * HL - 0.2 * HM - 0.6 * HN
 - 0.2 * HO + 0.2 * HP - 0.2 * HQ - 0.2 * HS - 0.2 * IJ - 0.2 * IK - 0.2 * IL
 + 0.2 * IM - 0.2 * IN + 0.2 * IO + 0.2 * IP + 0.2 * IQ - 0.2 * IS - 0.6 * JK
 - 0.2 * JL + 0.2 * JM - 0.2 * JN - 0.2 * JO + 0.2 * JP + 0.2 * JQ + 0.2 * JS
 + 0.2 * KL + 0.2 * KM + 0.2 * KN - 0.2 * KO - 0.2 * KP - 0.2 * KQ + 0.2 * KS
 - 0.2 * LM - 0.2 * LN - 0.2 * LO + 0.2 * LP + 0.2 * LQ - 0.2 * LS + 0.2 * MN
 - 0.6 * MO + 0.2 * MP - 0.2 * MQ + 0.2 * MS + 0.2 * NO - 0.2 * NP - 0.2 * NQ
 - 0.2 * NS + 0.2 * OP - 0.2 * OQ - 0.2 * OS - 0.2 * PQ - 0.2 * PS - 0.2 * QS
[S] = S - 0.2 * AB + 0.2 * AC + 0.2 * AD + 0.2 * AE - 0.2 * AF + 0.2 * AG
 - 0.6 * AH - 0.2 * AI - 0.2 * AJ + 0.2 * AK + 0.2 * AL - 0.2 * AM + 0.2 * AN
 - 0.2 * AO - 0.2 * AP - 0.2 * AQ - 0.2 * AR - 0.2 * BC - 0.2 * BD - 0.6 * BE
 - 0.2 * BF + 0.2 * BG - 0.2 * BH - 0.2 * BI - 0.2 * BJ + 0.2 * BK + 0.2 * BL
 + 0.2 * BM - 0.2 * BN + 0.2 * BO + 0.2 * BP - 0.2 * BQ + 0.2 * BR - 0.2 * CD
 + 0.2 * CE - 0.2 * CF + 0.2 * CG + 0.2 * CH - 0.2 * CI + 0.2 * CJ - 0.2 * CK
```

```
- 0.2 * CL + 0.2 * CM - 0.2 * CN - 0.2 * CO - 0.2 * CP - 0.6 * CQ + 0.2 * CR
- 0.2 * DE + 0.2 * DF - 0.2 * DG - 0.2 * DH + 0.2 * DI + 0.2 * DJ - 0.2 * DK
+ 0.2 * DL - 0.6 * DM - 0.2 * DN - 0.2 * DO + 0.2 * DP - 0.2 * DQ + 0.2 * DR
+ 0.2 * EF - 0.2 * EG - 0.2 * EH - 0.2 * EI - 0.2 * EJ - 0.2 * EK - 0.2 * EL
+ 0.2 * EM - 0.2 * EN + 0.2 * EO + 0.2 * EP + 0.2 * EQ - 0.2 * ER - 0.2 * FG
+ 0.2 * FH + 0.2 * FI - 0.6 * FJ + 0.2 * FK - 0.2 * FL - 0.2 * FM - 0.2 * FN
+ 0.2 * FO - 0.2 * FP - 0.2 * FQ + 0.2 * FR + 0.2 * GH + 0.2 * GI - 0.2 * GJ
+ 0.2 * GK - 0.2 * GL - 0.2 * GM - 0.2 * GN - 0.2 * GO + 0.2 * GP - 0.2 * GQ
- 0.6 * GR + 0.2 * HI + 0.2 * HJ - 0.2 * HK - 0.2 * HL - 0.2 * HM + 0.2 * HN
+ 0.2 * HO - 0.2 * HP - 0.2 * HQ - 0.2 * HR - 0.2 * IJ - 0.2 * IK + 0.2 * IL
+ 0.2 * IM - 0.2 * IN - 0.6 * IO - 0.2 * IP + 0.2 * IQ - 0.2 * IR - 0.2 * JK
- 0.2 * JL - 0.2 * JM + 0.2 * JN - 0.2 * JO + 0.2 * JP + 0.2 * JQ + 0.2 * JR
- 0.6 * KL - 0.2 * KM + 0.2 * KN - 0.2 * KO - 0.2 * KP + 0.2 * KQ + 0.2 * KR
+ 0.2 * LM + 0.2 * LN + 0.2 * LO - 0.2 * LP - 0.2 * LQ - 0.2 * LR - 0.2 * MN
- 0.2 * MO - 0.2 * MP + 0.2 * MQ + 0.2 * MR + 0.2 * NO - 0.6 * NP + 0.2 * NQ
- 0.2 * NR + 0.2 * OP - 0.2 * OQ - 0.2 * OR + 0.2 * PQ - 0.2 * PR - 0.2 * QR
```

(b) Show that for the 20-run Plackett-Burman in Table P8.10, the weights (or correlations) that multiply the two-factor interactions in each alias chain are either -0.2, 0.2, or -0.6. Of the 153 interactions that are aliased with each main effect, 144 have weights of -0.2 or 0.2, while nine interactions have weights of -0.6

See the Design Expert output shown above.

(c) Verify that the five largest main effects are S, G, R, I and J.



Design Expert Output

Response	**2**	**Response rate**				
ANOVA for selected factorial model						
Analysis of variance table [Partial sum of squares - Type III]						
	Sum of		**Mean**	**F**	**p-value**	
Source	**Squares**	**df**	**Square**	**Value**	**Prob > F**	
Model	6.36	5	1.27	35.43	< 0.0001	significant
G-Sticker	*1.55*	*1*	*1.55*	*43.03*	*< 0.0001*	
I-Copy Message	*0.44*	*1*	*0.44*	*12.20*	*0.0036*	
J-Letter headline	*0.18*	*1*	*0.18*	*5.13*	*0.0399*	
R-2nd buckslip	*0.46*	*1*	*0.46*	*12.86*	*0.0030*	
S-Interest rate	*3.73*	*1*	*3.73*	*103.91*	*< 0.0001*	
Residual	0.50	14	0.036			
Cor Total	6.87	19				

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	1.30	1	0.042	1.21	1.39	
G-Sticker	-0.28	1	0.042	-0.37	-0.19	1.00
J-Copy Message	0.15	1	0.042	0.057	0.24	1.00
K-Letter headline	-0.096	1	0.042	-0.19	-5.106E-003	1.00
S-2nd buckslip	-0.15	1	0.042	-0.24	-0.061	1.00
T-Interest rate	-0.43	1	0.042	-0.52	-0.34	1.00

(d) Factor S (interest rate) and G (presence of a sticker) are by far the largest main effects. The correlation between the main effect of R (2^{nd} buckslip) and the SG interaction is -0.6. This means that a significant SG interaction would bias the estimate of the main effect of R by -0.6 times the value of the interaction. This suggests that it may not be the main effect of factor R that is important, but the two factor interaction between S and G.

(e) Since this design projects into a full factorial in any three factors, obtain the projection in factors S, G and R and verify that it is a full factorial with some runs replicated. Fit a full factorial model involving all three of these factors and the interactions (you will need to use a regression program to do this). Show that S, G and the SG interaction are significant.

Test Cell	G	R	S	GR	GS	RS	GRS	run	Orders	Resp Rate
1	1	1	-1	1	-1	-1	-1	gr	52	1.04%
2	1	1	1	1	1	1	1	grs	38	0.76%
3	1	-1	1	-1	1	-1	-1	gs	42	0.84%
4	-1	-1	-1	1	1	1	-1	(1)	134	2.68%
5	-1	-1	-1	1	1	1	-1	(1)	104	2.08%
6	1	-1	-1	-1	-1	1	1	g	60	1.20%
7	1	1	-1	1	-1	-1	-1	gr	61	1.22%
8	-1	-1	1	1	-1	-1	1	s	68	1.36%
9	1	1	-1	1	-1	-1	-1	gr	57	1.14%
10	1	-1	1	-1	1	-1	-1	gs	30	0.60%
11	-1	1	-1	-1	1	-1	1	r	108	2.16%
12	-1	1	1	-1	-1	1	-1	rs	39	0.78%
13	-1	1	1	-1	-1	1	-1	rs	40	0.80%
14	-1	1	1	-1	-1	1	-1	rs	49	0.98%
15	1	-1	1	-1	1	-1	-1	gs	37	0.74%
16	-1	-1	-1	1	1	1	-1	(1)	99	1.98%
17	1	1	-1	1	-1	-1	-1	gr	86	1.72%
18	-1	1	1	-1	-1	1	-1	rs	43	0.86%
19	1	-1	1	-1	1	-1	-1	gs	47	0.94%
20	-1	-1	-1	-1	-1	-1	-1	gr	104	2.08%

The projected design is shown in the above table, along with the terms to estimate the interaction. Note that all 8 corners of the 2^3 are represented at least once. The analysis shown below shows that the significant factors are *G* and *S* (*A* and *C*). *R* (*B*) is not significant. The *GS* (*AC*) interaction has a higher *t*-value than the *R* (*C*) factor that it was confounded with in the original Plackett-Burman design.

Design Expert Output

Response 2 Response rate
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	6.20	7	0.89	15.87	< 0.0001	significant
A-G - Sticker	*1.31*	*1*	*1.31*	*23.50*	*0.0004*	
B-R - 2nd buckslip	*0.048*	*1*	*0.048*	*0.86*	*0.3718*	
C-S - Interest rate	*1.91*	*1*	*1.91*	*34.23*	*< 0.0001*	
AB	*0.074*	*1*	*0.074*	*1.33*	*0.2706*	
AC	*0.29*	*1*	*0.29*	*5.25*	*0.0408*	
BC	*0.063*	*1*	*0.063*	*1.12*	*0.3099*	
ABC	*0.026*	*1*	*0.026*	*0.46*	*0.5084*	
Pure Error	0.67	12	0.056			
Cor Total	6.87	19				

Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	1.33	1	0.066	1.18	1.47	
A-G - Sticker	-0.32	1	0.066	-0.46	-0.18	1.56
B-R - 2nd buckslip	-0.061	1	0.066	-0.21	0.083	1.56
C-S - Interest rate	-0.39	1	0.066	-0.53	-0.24	1.56
AB	0.076	1	0.066	-0.068	0.22	1.56
AC	0.15	1	0.066	7.414E-003	0.30	1.56
BC	-0.070	1	0.066	-0.21	0.074	1.56
ABC	0.045	1	0.066	-0.099	0.19	1.00

8.39. Consider the following experiment:

Run	Treatment Combination	A	B	C	D=AB	E=-AC
1	d	-1	-1	-1	1	-1
2	ae	1	-1	-1	-1	1
3	b	-1	1	-1	-1	-1
4	abde	1	1	-1	1	1
5	cde	-1	-1	1	1	1
6	ac	1	-1	1	-1	-1
7	bce	-1	1	1	-1	1
8	abcd	1	1	1	1	-1

Answer the following questions about this experiment:

(a) How many factors did this experiment investigate? 5
(b) How many factors are in the basic design? 3
(c) Assume that the factors in the experiment are represented by the initial letters of the alphabet (i.e. A, B, etc.), what are the design generators for the factors beyond the basic design? D = AB and E = -AC, therefore I = ABD - ACE
(d) Is this design a principal fraction? No
(e) What is the complete defining relation? I = ABD –ACE – BDCE
(f) What is the resolution of this design? III

8.40. Consider the following experiment:

Run	Treatment Combination		A	B	C	D=ABC	y
1	(1)		-1	-1	-1	-1	8
2	ad		1	-1	-1	1	10
3	bd		-1	1	-1	1	12
4	ab		1	1	-1	-1	7
5	cd		-1	-1	1	1	13
6	ac		1	-1	1	-1	6
7	bc		-1	1	1	-1	5
8	abcd		1	1	1	1	11
		Avg (+)	8.5	4.5	8.75	11.5	
		Avg (-)	9.5	9.25	9.25	26	
		Effect	-0.5	-2.375	-0.25	-7.25	

(a) How many factors did this experiment investigate? 4
(b) What is the resolution of this design? IV
(c) Calculate the estimates of the effects. See table above.
(d) What is the complete defining relation? I = ABCD

8.41. An unreplicated 2^{5-1} fractional factorial experiment with four center points has been run in a chemical process. The response variable is molecular weight. The experimenter has used the following factors:

Factor	Natural levels	Coded levels (x's)
A – time	20, 40 (minutes)	-1, 1
B – temperature	160, 180 (deg C)	-1, 1
C – concentration	30, 60 (percent)	-1, 1
D – stirring rate	100, 150 (rpm)	-1, 1
E – catalyst type	1, 2 (type)	-1, 1

Suppose that the prediction equation from this experiment is $\hat{y} = 10 + 3x_1 + 2x_2 - 1x_1x_2$. What is the predicted response at A = 30, B = 165, C = 50, D = 135, and E = 1?

A = 3, is A = 0 in coded units, B = 165 is B = -0.5 in coded units. These are the only two factors in the prediction equation.

$$\hat{y} = 10 + 3(0) + 2(-0.5) - 1(0)(-0.5) = 9$$

8.42. An unreplicated 2^{4-1} fractional factorial experiment with four center points has been run. The experimenter has used the following factors:

Factor	Natural levels	Coded levels (x's)
A – time	10, 50 (minutes)	-1, 1
B – temperature	200, 300 (deg C)	-1, 1
C – concentration	70, 90 (percent)	-1, 1
D – pressure	260, 300 (psi)	-1, 1

(a) Suppose that the average of the 8 factorial design points is 100 and the average of the center points is 120, what is the sum of squares for pure quadratic curvature?

$$SS_{\text{Pure quadratic}} = \frac{n_F n_C\left(\bar{y}_F - \bar{y}_C\right)^2}{n_F + n_C} = \frac{8*4\left(100-120\right)^2}{8+4} = 1066.67$$

(b) Suppose that the prediction equation that results from this experiment is $\hat{y} = 50 + 5x_1 + 2x_2 - 2x_1x_2$. Find the predicted response at A = 20, B = 250, C = 80 and D = 275.

A = 20, is A = -0.5 in coded units, B = 250 is B = 0 in coded units. These are the only two factors in the prediction equation.

$$\hat{y} = 50 + 5(-0.5) + 2(0) - 2(-0.50)(0) = 47.5$$

8.43. An unreplicated 2^{4-1} fractional factorial experiment has been run. The experimenter has used the following factors:

Factor	Natural levels	Coded levels (x's)
A	20, 50	-1, 1
B	200, 280	-1, 1
C	50, 100	-1, 1
D	150, 200	-1, 1

(a) Suppose that this design has four center points that average 100. The average of the 8 factorial design points is 95. What is the sum of squares for pure quadratic curvature?

$$SS_{\text{Pure quadratic}} = \frac{n_F n_C\left(\bar{y}_F - \bar{y}_C\right)^2}{n_F + n_C} = \frac{8*4\left(95-100\right)^2}{8+4} = 66.67$$

(b) Suppose that the prediction equation that results from this experiment is $\hat{y} = 100 - 2x_1 + 10x_2 - 4x_1x_2$. Find the predicted response at A = 41, B = 280, C = 60 and D = 185.

A = 41, is A = 0.4 in coded units, B = 280 is B = 1 in coded units. These are the only two factors in the prediction equation.

$$\hat{y} = 100 - 2(0.4) + 10(1) - 4(0.4)(10) = 93.2$$

8.44. A 2^{6-2} factorial experiment with three replicates has been run in a pharmaceutical drug manufacturing process. The experimenter has used the following factors:

Factor	Natural levels	Coded levels (x's)
A	50, 100	-1, 1
B	20, 60	-1, 1
C	10, 30	-1, 1
D	12, 18	-1, 1
E	15, 30	-1, 1
F	60, 100	-1, 1

(a) If two main effects and one two-factor interaction are included in the final model, how many degrees of freedom for error will be available? A 2^{6-2} factorial experiment has 16 runs. We are estimating four model parameters – the intercept, two main effects and one two-factor interactions. This leaves 12 degrees of freedom for error. This assumes a non-hierarchical model.

(b) Suppose that the significant factors are A, C, AB and AC. What other effects need to be included to obtain a hierarchical model? The hierarchical model would include A, B, C, AB and AC.

8.45. Consider the following design:

Run	A	B	C	D	E	y	block = ABC
1	-1	-1	-1	-1	-1	63	-1
2	1	-1	-1	-1	1	21	1
3	-1	1	-1	-1	1	36	1
4	1	1	-1	-1	-1	99	-1
5	-1	-1	1	-1	1	24	1
6	1	-1	1	-1	-1	66	-1
7	-1	1	1	-1	-1	71	-1
8	1	1	1	-1	1	54	1
9	-1	-1	-1	1	-1	23	-1
10	1	-1	-1	1	1	74	1
11	-1	1	-1	1	1	80	1
12	1	1	-1	1	-1	33	-1
13	-1	-1	1	1	1	63	1
14	1	-1	1	1	-1	21	-1
15	-1	1	1	1	-1	44	-1
16	1	1	1	1	1	96	1
Avg(+)	58	64.125	54.875	54.25	56		
Avg(-)	50.5	44.375	53.625	54.25	52.5		
Effect	3.75	9.875	0.625	0	1.75		

(a) What is the generator for column E? E = ABC
(b) If ABC is confounded with blocks, run 1 above goes in the ___ block. Answer either + or -. ABC goes in to the – block.
(c) What is the resolution of this design? IV
(d) Find the estimates of the main effects and their aliases. See Table.

8.46. Consider the following design:

Run	A	B	C	D	E	y	block=ABE
1	-1	-1	-1	-1	-1	65	-1
2	1	-1	-1	-1	1	25	-1
3	-1	1	-1	-1	1	30	-1
4	1	1	-1	-1	-1	89	-1
5	-1	-1	1	-1	1	25	1
6	1	-1	1	-1	-1	60	1
7	-1	1	1	-1	-1	70	1
8	1	1	1	-1	1	50	1
9	-1	-1	-1	1	1	20	1
10	1	-1	-1	1	-1	70	1
11	-1	1	-1	1	-1	80	1
12	1	1	-1	1	1	30	1

13	-1	-1	1	1	-1	60	-1
14	1	-1	1	1	1	20	-1
15	-1	1	1	1	1	40	-1
16	1	1	1	1	-1	90	-1
Avg(+)	54.25	59.875	51.875	51.25	30		
Avg(-)	48.75	43.125	51.125	51.75	73		
Effect	2.75	8.375	0.375	-0.25	-21.5		

(a) What is the generator for column E? E = -ABCD
(b) If ABE is confounded with blocks, run 16 above goes in the ___ block. Answer either + or -. Run 16 is in the – block.
(c) What is the resolution of this design? V
(d) Find the estimates of the main effects and their aliases See Table.

8.47. Consider the following design:

Run	A	B	C	D=-ABC	E=-BC	Block = AB
1	-1	-1	-1	1	-1	1
2	1	-1	-1	-1	-1	-1
3	-1	1	-1	-1	1	-1
4	1	1	-1	1	1	1
5	-1	-1	1	-1	1	1
6	1	-1	1	1	1	-1
7	-1	1	1	1	-1	-1
8	1	1	1	-1	-1	1

(a) What is the generator for column D? D = -ABC
(b) What is the generator for column E? E = -BC
(c) If this design were run in two blocks with the AB interaction confounded with blocks, the run *d* would be in the block where the sign of AB is ____? Answer either + or -. Run *d*, is the combination (-1, -1, -1, 1, -1), it is run 1 and is in the + block.

8.48. Consider the following design:

Run	A	B	C	D	E
1	-1	-1	-1	1	1
2	1	-1	-1	-1	1
3	-1	1	-1	-1	-1
4	1	1	-1	1	-1
5	-1	-1	1	-1	-1
6	1	-1	1	1	-1
7	-1	1	1	1	1
8	1	1	1	-1	1

9	1	1	1	-1	-1
10	-1	1	1	1	-1
11	1	-1	1	1	1
12	-1	-1	1	-1	1
13	1	1	-1	1	1
14	-1	1	-1	-1	1
15	1	-1	-1	-1	-1
16	-1	-1	-1	1	-1

(a) What is the generator for column D? D = -ABC
(b) What is the generator for column E? E = BC
(c) If this design were folded over, what is the resolution of the combined design? IV

8.49. In an article in *Quality Engineering* ("An Application of Fractional Factorial Experimental Designs," 1988, Vol. 1 pp. 19-23) M.B. Kilgo describes an experiment to determine the effect of CO_2 pressure (*A*), CO_2 temperature (*B*), peanut moisture (*C*), CO_2 flow rate (*D*), and peanut particle size (*E*) on the total yield of oil per batch of peanuts (*y*). The levels she used for these factors are shown in Table P8.11. She conducted the 16-run fractional factorial experiment shown in Table P8.12.

Table P8.11

Coded Level	*A* Pressure (bar)	*B* Temp (C)	*C* Moisture (% by weight)	*D* Flow (liters/min)	*E* Particle Size (mm)
-1	415	25	5	40	1.28
1	550	95	15	60	4.05

Table P8.12

	A	*B*	*C*	*D*	*E*	y
1	415	25	5	40	1.28	63
2	550	25	5	40	4.05	21
3	415	95	5	40	4.05	36
4	550	95	5	40	1.28	99
5	415	25	15	40	4.05	24
6	550	25	15	40	1.28	66
7	415	95	15	40	1.28	71
8	550	95	15	40	4.05	54
9	415	25	5	60	4.05	23
10	550	25	5	60	1.28	74
11	415	95	5	60	1.28	80
12	550	95	5	60	4.05	33
13	415	25	15	60	1.28	63
14	550	25	15	60	4.05	21
15	415	95	15	60	4.05	44
16	550	95	15	60	1.28	96

(a) What type of design has been used? Identify the defining relation and the alias relationships.

A 2_V^{5-1}, 16-run design, with I= *-ABCDE*.

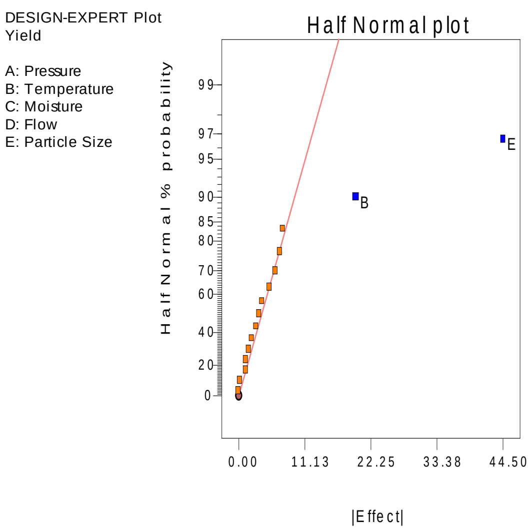
A(-ABCDE)=	*-BCDE*	*A=*	*-BCDE*
B(-ABCDE)=	*-ACDE*	*B=*	*-ACDE*
C(-ABCDE)=	*-ABDE*	*C=*	*-ABDE*

D(-ABCDE)=	*-ABCE*	*D=*	*-ABCE*
E(-ABCDE)=	*-ABCD*	*E=*	*-ABCD*
AB(-ABCDE)=	*-CDE*	*AB=*	*-CDE*
AC(-ABCDE)=	*-BDE*	*AC=*	*-BDE*
AD(-ABCDE)=	*-BCE*	*AD=*	*-BCE*
AE(-ABCDE)=	*-BCD*	*AE=*	*-BCD*
BC(-ABCDE)=	*-ADE*	*BC=*	*-ADE*
BD(-ABCDE)=	*-ACE*	*BD=*	*-ACE*
BE(-ABCDE)=	*-ACD*	*BE=*	*-ACD*
CD(-ABCDE)=	*-ABE*	*CD=*	*-ABE*
CE(-ABCDE)=	*-ABD*	*CE=*	*-ABD*
DE(-ABCDE)=	*-ABC*	*DE=*	*-ABC*

(b) Estimate the factor effects and use a normal probability plot to tentatively identify the important factors.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	7.5	225	2.17119
Model	B	19.75	1560.25	15.056
Error	C	1.25	6.25	0.0603107
Error	D	0	0	0
Model	E	44.5	7921	76.4354
Error	AB	5.25	110.25	1.06388
Error	AC	1.25	6.25	0.0603107
Error	AD	-4	64	0.617582
Error	AE	7	196	1.89134
Error	BC	3	36	0.34739
Error	BD	-1.75	12.25	0.118209
Error	BE	0.25	0.25	0.00241243
Error	CD	2.25	20.25	0.195407
Error	CE	-6.25	156.25	1.50777
Error	DE	3.5	49	0.472836
	Lenth's ME	11.5676		
	Lenth's SME	23.4839		



(c) Perform an appropriate statistical analysis to test the hypothesis that the factors identified in part above have a significant effect on the yield of peanut oil.

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	9481.25	2	4740.63	69.89	< 0.0001	significant
B	*1560.25*	*1*	*1560.25*	*23.00*	*0.0003*	
E	*7921.00*	*1*	*7921.00*	*116.78*	*< 0.0001*	
Residual	881.75	13	67.83			
Cor Total	10363.00	15				

The Model F-value of 69.89 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	8.24	R-Squared	0.9149
Mean	54.25	Adj R-Squared	0.9018
C.V.	15.18	Pred R-Squared	0.8711
PRESS	1335.67	Adeq Precision	18.017

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	54.25	1	2.06	49.80	58.70	
B-Temperature	9.88	1	2.06	5.43	14.32	1.00
E-Particle Size	22.25	1	2.06	17.80	26.70	1.00

(d) Fit a model that could be used to predict peanut oil yield in terms of the factors that you have identified as important.

Design Expert Output

Final Equation in Terms of Coded Factors:

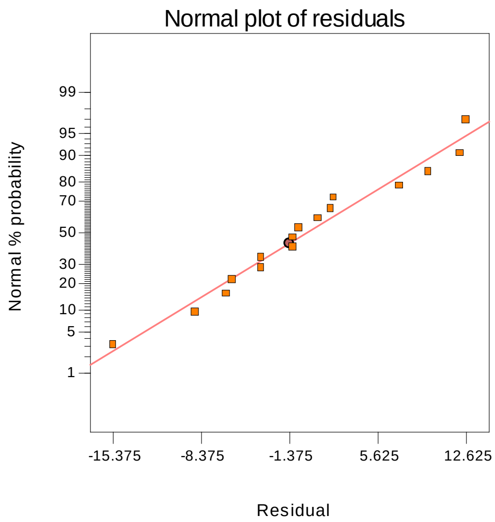
```
Yield    =
+54.25
+9.88     * B
+22.25    * E
```

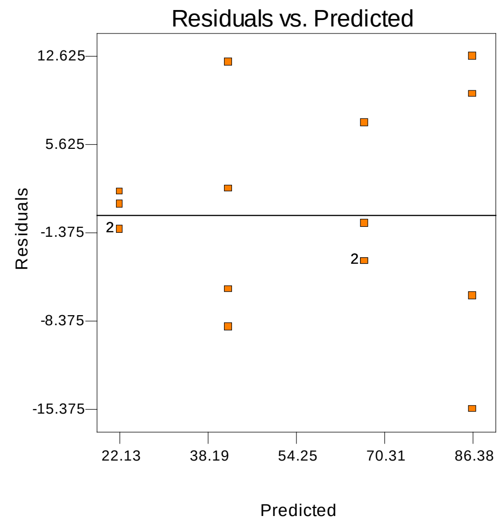
Final Equation in Terms of Actual Factors:

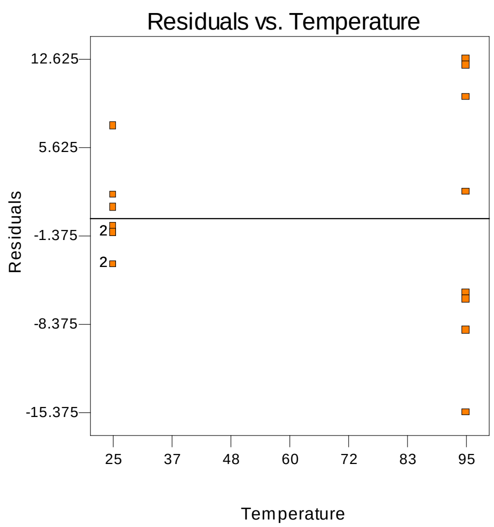
```
Yield    =
-5.49175
+0.28214    * Temperature
+16.06498   * Particle Size
```

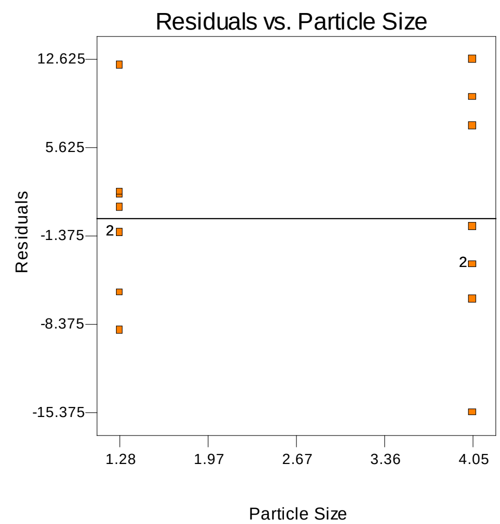
(e) Analyze the residuals from this experiment and comment on model adequacy.

The residual plots are satisfactory. There is a slight tendency for the variability of the residuals to increase with the predicted value of *y*.









8.50. A 16-run fractional factorial experiment in 10 factors on sand-casting of engine manifolds was conducted by engineers at the Essex Aluminum Plant of the Ford Motor Company and described in the article "Evaporative Cast Process 3.0 Liter Intake Manifold Poor Sandfill Study," by D. Becknell (*Fourth Symposium on Taguchi Methods*, American Supplier Institute, Dearborn, MI, 1986, pp. 120-130). The purpose was to determine which of 10 factors has an effect on the proportion of defective castings. The design and the resulting proportion of nondefective castings $\hat{p}$ observed on each run are shown in Table P8.13. This is a resolution III fraction with generators $E=CD$, $F=BD$, $G=BC$, $H=AC$, $J=AB$, and $K=ABC$. Assume that the number of castings made at each run in the design is 1000.

Table P8.13

Run	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	p	arcsin	F&T's Modification
1	-	-	-	-	+	+	+	+	+	-	0.958	1.364	1.363
2	+	-	-	-	+	+	+	-	-	+	1.000	1.571	1.555
3	-	+	-	-	+	-	-	+	-	+	0.977	1.419	1.417
4	+	+	-	-	+	-	-	-	+	-	0.775	1.077	1.076
5	-	-	+	-	-	+	-	-	+	+	0.958	1.364	1.363
6	+	-	+	-	-	+	-	+	-	-	0.958	1.364	1.363

7	-	+	+	-	-	-	+	-	-	-	0.813	1.124	1.123
8	+	+	+	-	-	-	+	+	+	+	0.906	1.259	1.259
9	-	-	-	+	-	-	+	+	+	-	0.679	0.969	0.968
10	+	-	-	+	-	-	+	-	-	+	0.781	1.081	1.083
11	-	+	-	+	-	+	-	+	-	+	1.000	1.571	1.556
12	+	+	-	+	-	+	-	-	+	-	0.896	1.241	1.242
13	-	-	+	+	+	-	-	-	+	+	0.958	1.364	1.363
14	+	-	+	+	+	-	-	+	-	-	0.818	1.130	1.130
15	-	+	+	+	+	+	+	-	-	-	0.841	1.161	1.160
16	+	+	+	+	+	+	+	+	+	+	0.955	1.357	1.356

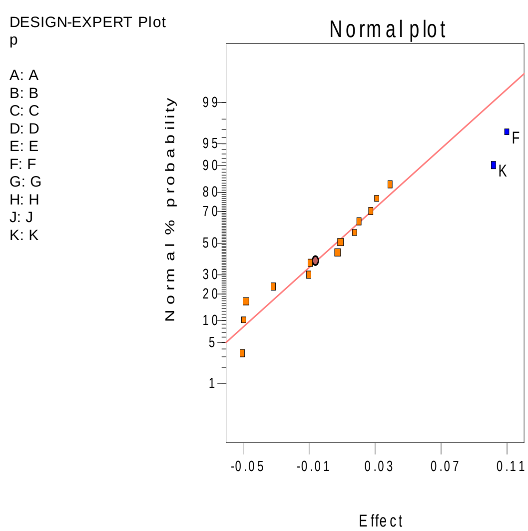
(a) Find the defining relation and the alias relationships in this design.

I=CDE=BDF=BCG=ACH=ABJ=ABCK=BCEF=BDEG=ADEH=ABCDEJ=ABDEK=CDFG=ABCDFH = ADFJ=ACDFK=ABGH=ACGJ=AGK=BCHJ=BHK=CjK

(b) Estimate the factor effects and use a normal probability plot to tentatively identify the important factors.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	-0.011875	0.000564063	0.409171
Error	B	0.006625	0.000175562	0.127353
Error	C	0.017625	0.00124256	0.901355
Error	D	-0.052125	0.0108681	7.88369
Error	E	0.036375	0.00529256	3.83923
Model	F	0.107375	0.0461176	33.4537
Error	G	-0.050875	0.0103531	7.51011
Error	H	0.028625	0.00327756	2.37754
Error	J	-0.012875	0.000663062	0.480986
Model	K	0.099625	0.0397006	28.7988
Error	AB	Aliased		
Error	AC	Aliased		
Error	AD	0.004875	9.50625E-005	0.0689584
Error	AE	-0.034625	0.00479556	3.4787
Error	AF	0.024875	0.00247506	1.79541
Error	BE	-0.053125	0.0112891	8.18909
Error	DK	0.015375	0.000945563	0.685911
	Lenth's ME	0.103145		
	Lenth's SME	0.209399		



(c) Fit an appropriate model using the factors identified in part (b) above.

Design Expert Output

Response: p
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	0.086	2	0.043	10.72	0.0018	significant
F	*0.046*	*1*	*0.046*	*11.52*	*0.0048*	
K	*0.040*	*1*	*0.040*	*9.92*	*0.0077*	
Residual	0.052	13	4.003E-003			
Cor Total	0.14	15				

The Model F-value of 10.72 implies the model is significant. There is only a 0.18% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.063	R-Squared	0.6225
Mean	0.89	Adj R-Squared	0.5645
C.V.	7.09	Pred R-Squared	0.4282
PRESS	0.079	Adeq Precision	7.556

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.89	1	0.016	0.86	0.93	
F-F	0.054	1	0.016	0.020	0.088	1.00
K-K	0.050	1	0.016	0.016	0.084	1.00

Final Equation in Terms of Coded Factors:

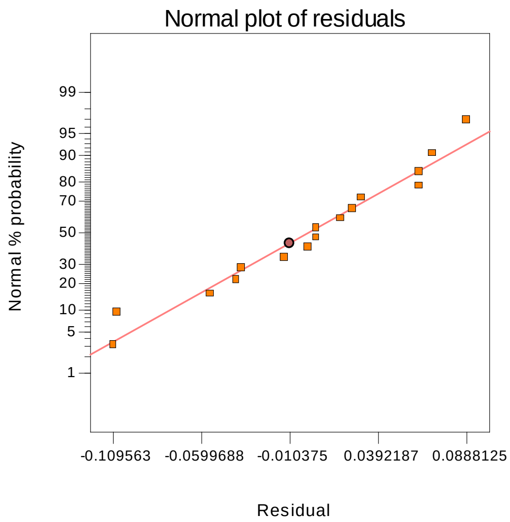
p =
+0.89
+0.054 * F
+0.050 * K

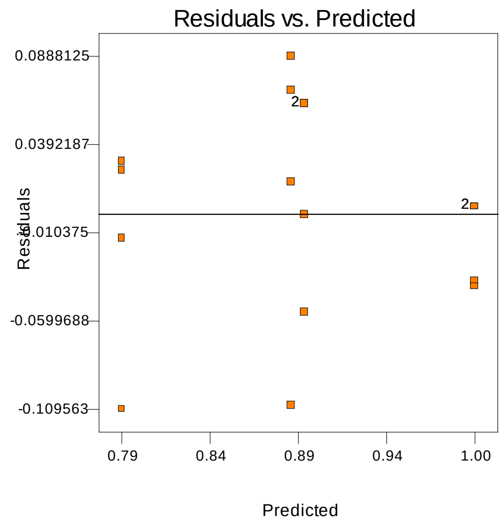
Final Equation in Terms of Actual Factors:

p =
+0.89206
+0.053688 * F
+0.049812 * K

(d) Plot the residuals from this model versus the predicted proportion of nondefective castings. Also prepare a normal probability plot of the residuals. Comment on the adequacy of these plots.

The residual versus predicted plot identifies an inequality of variances. This is likely caused by the response variable being a proportion. A transformation could be used to correct this.



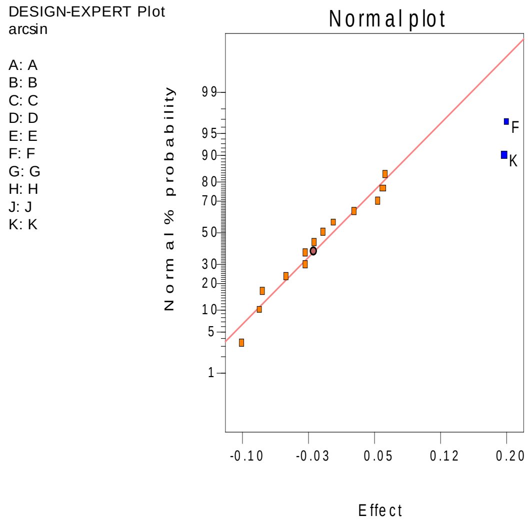


(e) In part (d) you should have noticed an indication that the variance of the response is not constant (considering that the response is a proportion, you should have expected this). The previous table also shows a transformation on $\hat{p}$, the arcsin square root, that is a widely used variance stabilizing transformation for proportion data (refer to the discussion of variance stabilizing transformations is Chapter 3). Repeat parts (a) through (d) above using the transformed response and comment on your results. Specifically, are the residuals plots improved?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	-0.032	0.004096	0.884531
Error	B	0.00025	2.5E-007	5.39875E-005
Error	C	-0.02125	0.00180625	0.39006
Error	D	-0.0835	0.027889	6.02263
Error	E	0.05875	0.0138062	2.98146
Model	F	0.19625	0.154056	33.2685
Error	G	-0.0805	0.025921	5.59764
Error	H	0.05625	0.0126562	2.73312
Error	J	-0.05325	0.0113422	2.44936
Model	K	0.1945	0.151321	32.6778
Error	AD	-0.032	0.004096	0.884531
Error	AF	0.05025	0.0101003	2.18115
Error	BE	-0.104	0.043264	9.34286
Error	DH	-0.01125	0.00050625	0.109325
Error	DK	0.0235	0.002209	0.477034
	Lenth's ME	0.205325		
	Lenth's SME	0.41684		

As with the original analysis, factors F and K remain significant with a slight increase with the R^2.



Design Expert Output

Response: **arcsin**

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	0.31	2	0.15	12.59	0.0009	significant
F	*0.15*	*1*	*0.15*	*12.70*	*0.0035*	
K	*0.15*	*1*	*0.15*	*12.47*	*0.0037*	
Residual	0.16	13	0.012			
Cor Total	0.46	15				

The Model F-value of 12.59 implies the model is significant. There is only a 0.09% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.11	R-Squared	0.6595
Mean	1.28	Adj R-Squared	0.6071
C.V.	8.63	Pred R-Squared	0.4842
PRESS	0.24	Adeq Precision	8.193

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	1.28	1	0.028	1.22	1.34	
F-F	0.098	1	0.028	0.039	0.16	1.00
K-K	0.097	1	0.028	0.038	0.16	1.00

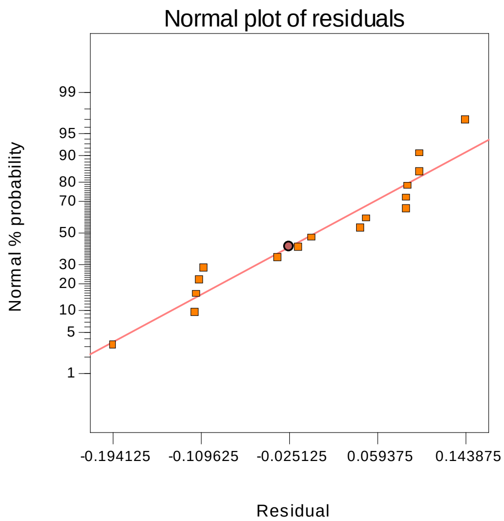
Final Equation in Terms of Coded Factors:

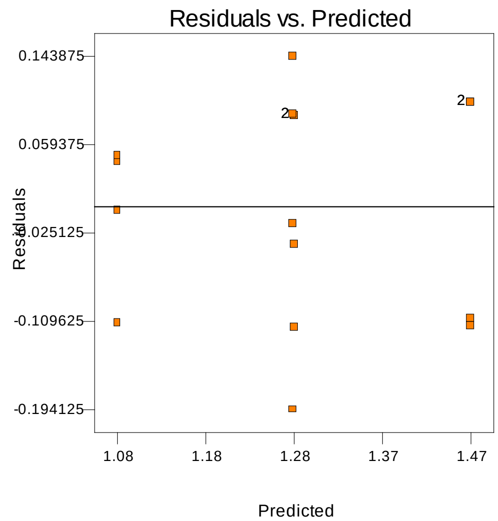
arcsin =
+1.28
+0.098 * F
+0.097 * K

Final Equation in Terms of Actual Factors:

arcsin =
+1.27600
+0.098125 * F
+0.097250 * K

The inequality of variance has improved; however, there remain hints of inequality in the residuals versus predicted plot and the normal probability plot now appears to be irregular.





(f) There is a modification to the arcsin square root transformation, proposed by Freeman and Tukey ("Transformations Related to the Angular and the Square Root," *Annals of Mathematical Statistics*, Vol. 21, 1950, pp. 607-611) that improves its performance in the tails. F&T's modification is:

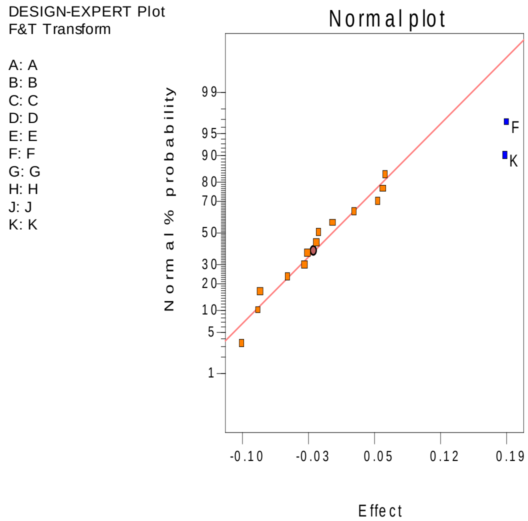
$$\frac{1}{2}\left[\arcsin\sqrt{\frac{n\hat{p}}{(n+1)}}+\arcsin\sqrt{\frac{(n\hat{p}+1)}{(n+1)}}\right]$$

Rework parts (a) through (d) using this transformation and comment on the results. (For an interesting discussion and analysis of this experiment, refer to "Analysis of Factorial Experiments with Defects or Defectives as the Response," by S. Bisgaard and H.T. Fuller, *Quality Engineering*, Vol. 7, 1994-5, pp. 429-443.)

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	-0.031125	0.00387506	0.871894
Error	B	0.000125	6.25E-008	1.40626E-005
Error	C	-0.017875	0.00127806	0.287566
Error	D	-0.082625	0.0273076	6.14424
Error	E	0.057875	0.0133981	3.01458
Model	F	0.192375	0.148033	33.3075
Error	G	-0.080375	0.0258406	5.81416
Error	H	0.055875	0.0124881	2.80983
Error	J	-0.049625	0.00985056	2.21639
Model	K	0.190875	0.145733	32.7901
Error	AD	-0.027875	0.00310806	0.699318
Error	AF	0.049625	0.00985056	2.21639
Error	BE	-0.100625	0.0405016	9.1129
Error	DH	-0.015375	0.000945563	0.212753
Error	DK	0.023625	0.00223256	0.502329
	Lenth's ME	0.191348		
	Lenth's SME	0.388464		

As with the prior analysis, factors *F* and *K* remain significant.



Design Expert Output

Response: F&T Transform
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	0.15	12.67	0.0009	significant
F	*0.15*	*1*	*0.15*	*12.77*	*0.0034*	
K	*0.15*	*1*	*0.15*	*12.57*	*0.0036*	
Residual	0.15	13	0.012			
Cor Total	0.44	15				

The Model F-value of 12.67 implies the model is significant. There is only a 0.09% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.11	R-Squared	0.6610
Mean	1.27	Adj R-Squared	0.6088
C.V.	8.45	Pred R-Squared	0.4864
PRESS	0.23	Adeq Precision	8.221

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	1.27	1	0.027	1.22	1.33	
F-F	0.096	1	0.027	0.038	0.15	1.00
K-K	0.095	1	0.027	0.037	0.15	1.00

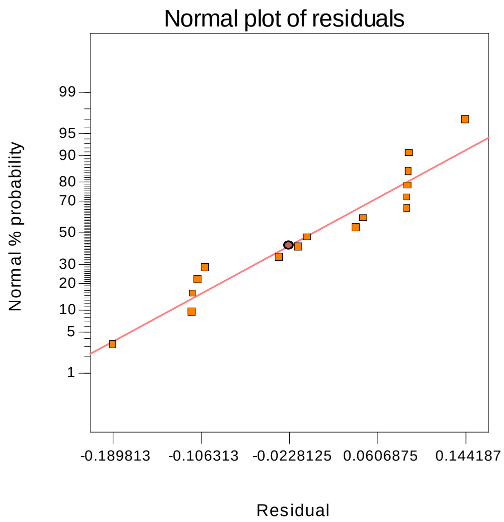
Final Equation in Terms of Coded Factors:

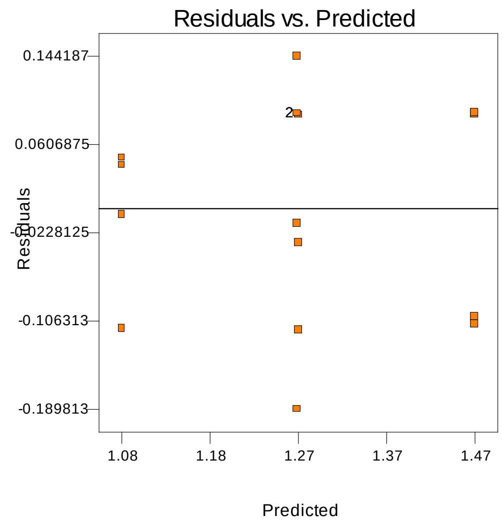
F&T Transform =
+1.27
+0.096 * F
+0.095 * K

Final Equation in Terms of Actual Factors:

F&T Transform =
+1.27356
+0.096188 * F
+0.095437 * K

The residual plots appear as they did with the arcsin square root transformation.





8.51. A 16-run fractional factorial experiment in nine factors was conducted by Chrysler Motors Engineering and described in the article "Sheet Molded Compound Process Improvement," by P.I. Hsieh and D.E. Goodwin (*Fourth Symposium on Taguchi Methods*, American Supplier Institute, Dearborn, MI, 1986, pp. 13-21). The purpose was to reduce the number of defects in the finish of sheet-molded grill opening panels. The design, and the resulting number of defects, c, observed on each run, is shown in Table P8.10. This is a resolution III fraction with generators $E=BD$, $F=BCD$, $G=AC$, $H=ACD$, and $J=AB$.

Table P8.10

Run	A	B	C	D	E	F	G	H	J	c	$\sqrt{c}$	F&T's Modification
1	-	-	-	-	+	-	+	-	+	56	7.48	7.52
2	+	-	-	-	+	-	-	+	-	17	4.12	4.18
3	-	+	-	-	-	+	+	-	-	2	1.41	1.57
4	+	+	-	-	-	+	-	+	+	4	2.00	2.12
5	-	-	+	-	+	+	-	+	+	3	1.73	1.87
6	+	-	+	-	+	+	+	-	-	4	2.00	2.12
7	-	+	+	-	-	-	-	+	-	50	7.07	7.12
8	+	+	+	-	-	-	+	-	+	2	1.41	1.57
9	-	-	-	+	-	+	+	+	+	1	1.00	1.21
10	+	-	-	+	-	+	-	-	-	0	0.00	0.50
11	-	+	-	+	+	-	+	+	-	3	1.73	1.87
12	+	+	-	+	+	-	-	-	+	12	3.46	3.54
13	-	-	+	+	-	-	-	-	+	3	1.73	1.87
14	+	-	+	+	-	-	+	+	-	4	2.00	2.12
15	-	+	+	+	+	+	-	-	-	0	0.00	0.50
16	+	+	+	+	+	+	+	+	+	0	0.00	0.50

(a) Find the defining relation and the alias relationships in this design.

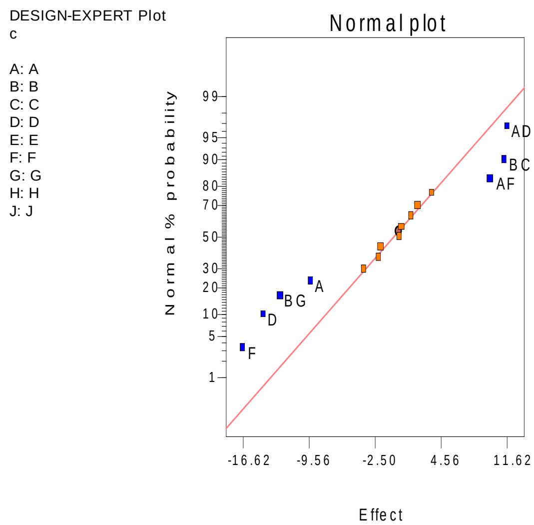
I = ABJ = ACG = BDE = CEF = DGH = FHJ = ABFH = ACDH = ADEJ = AEFG = BCDF = BCGJ = BEGH = CEHJ = DFGJ = ABCEH = ABDFG = ACDFJ = ADEFH = AEGHJ = BCDHJ = BCFGH = BEFGJ = CDEGJ = ABCDEG = ABCEFJ = ABDGHJ = ACFGHJ = BDEFHJ = CDEFGH = ABCDEFGHJ

(b) Estimate the factor effects and use a normal probability plot to tentatively identify the important factors.

The effects are shown below in the *Design Expert* output. The normal probability plot of effects identifies factors *A*, *D*, *F*, and interactions *AD*, *AF*, *BC*, *BG* as important.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	-9.375	351.562	7.75573
Model	B	-1.875	14.0625	0.310229
Model	C	-3.625	52.5625	1.15957
Model	D	-14.375	826.562	18.2346
Error	E	3.625	52.5625	1.15957
Model	F	-16.625	1105.56	24.3895
Model	G	-2.125	18.0625	0.398472
Error	H	0.375	0.5625	0.0124092
Error	J	0.125	0.0625	0.0013788
Model	AD	11.625	540.563	11.9252
Error	AE	2.125	18.0625	0.398472
Model	AF	9.875	390.063	8.60507
Error	AH	1.375	7.5625	0.166834
Model	BC	11.375	517.563	11.4178
Model	BG	-12.625	637.562	14.0651
	Lenth's ME	13.9775		
	Lenth's SME	28.3764		



(c) Fit an appropriate model using the factors identified in part (b) above.

The analysis of variance and corresponding model is shown below. Factors *B*, *C*, and *G* are included for hierarchal purposes.

Design Expert Output

Response: c

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	4454.13	10	445.41	28.26	0.0009	significant
A	*351.56*	*1*	*351.56*	*22.30*	*0.0052*	
B	*14.06*	*1*	*14.06*	*0.89*	*0.3883*	
C	*52.56*	*1*	*52.56*	*3.33*	*0.1274*	
D	*826.56*	*1*	*826.56*	*52.44*	*0.0008*	
F	*1105.56*	*1*	*1105.56*	*70.14*	*0.0004*	
G	*18.06*	*1*	*18.06*	*1.15*	*0.3333*	
AD	*540.56*	*1*	*540.56*	*34.29*	*0.0021*	

AF	*390.06*	*1*	*390.06*	*24.75*	*0.0042*
BC	*517.56*	*1*	*517.56*	*32.84*	*0.0023*
BG	*637.56*	*1*	*637.56*	*40.45*	*0.0014*
Residual	78.81	5	15.76		
Cor Total	4532.94	15			

The Model F-value of 28.26 implies the model is significant. There is only a 0.09% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	3.97	R-Squared	0.9826
Mean	10.06	Adj R-Squared	0.9478
C.V.	39.46	Pred R-Squared	0.8220
PRESS	807.04	Adeq Precision	17.771

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	10.06	1	0.99	7.51	12.61	
A-A	-4.69	1	0.99	-7.24	-2.14	1.00
B-B	-0.94	1	0.99	-3.49	1.61	1.00
C-C	-1.81	1	0.99	-4.36	0.74	1.00
D-D	-7.19	1	0.99	-9.74	-4.64	1.00
F-F	-8.31	1	0.99	-10.86	-5.76	1.00
G-G	-1.06	1	0.99	-3.61	1.49	1.00
AD	5.81	1	0.99	3.26	8.36	1.00
AF	4.94	1	0.99	2.39	7.49	1.00
BC	5.69	1	0.99	3.14	8.24	1.00
BG	-6.31	1	0.99	-8.86	-3.76	1.00

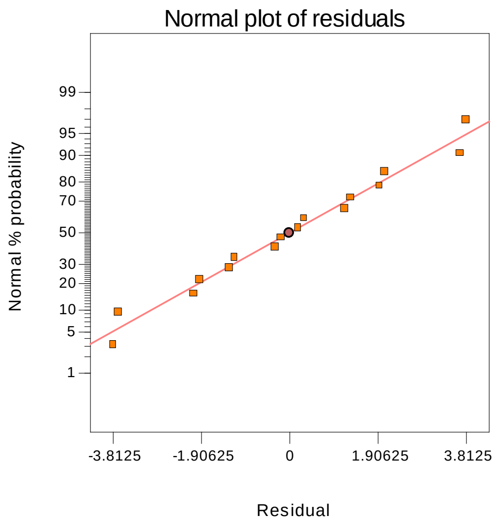
Final Equation in Terms of Coded Factors:

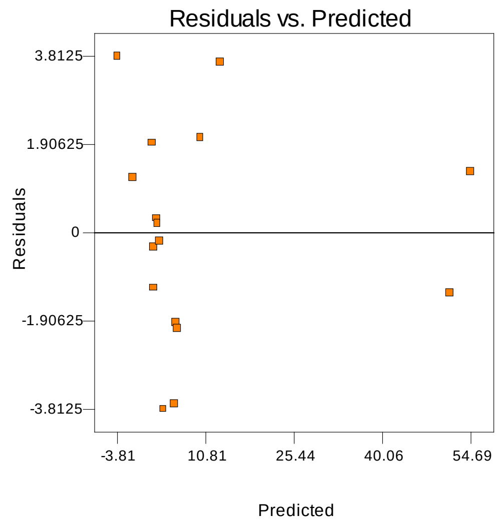
c	=
+10.06	
-4.69	* A
-0.94	* B
-1.81	* C
-7.19	* D
-8.31	* F
-1.06	* G
+5.81	* A * D
+4.94	* A * F
+5.69	* B * C
-6.31	* B * G

Final Equation in Terms of Actual Factors:

c	=
+10.06250	
-4.68750	* A
-0.93750	* B
-1.81250	* C
-7.18750	* D
-8.31250	* F
-1.06250	* G
+5.81250	* A * D
+4.93750	* A * F
+5.68750	* B * C
-6.31250	* B * G

(d) Plot the residuals from this model versus the predicted number of defects. Also, prepare a normal probability plot of the residuals. Comment on the adequacy of these plots.





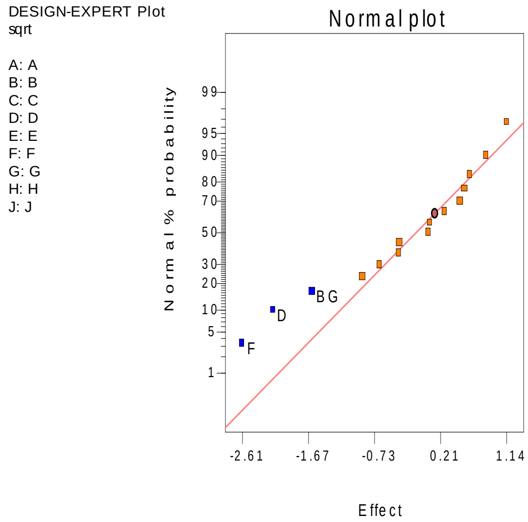
There is a significant problem with inequality of variance. This is likely caused by the response variable being a count. A transformation may be appropriate.

(e) In part (d) you should have noticed an indication that the variance of the response is not constant (considering that the response is a count, you should have expected this). The previous table also shows a transformation on c, the square root, that is a widely used variance stabilizing transformation for count data (refer to the discussion of variance stabilizing transformations in Chapter 3). Repeat parts (a) through (d) using the transformed response and comment on your results. Specifically, are the residual plots improved?

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	-0.895	3.2041	4.2936
Model	B	-0.3725	0.555025	0.743752
Error	C	-0.6575	1.72922	2.31722
Model	D	-2.1625	18.7056	25.0662
Error	E	0.4875	0.950625	1.27387
Model	F	-2.6075	27.1962	36.4439
Model	G	-0.385	0.5929	0.794506
Error	H	0.27	0.2916	0.390754
Error	J	0.06	0.0144	0.0192965
Error	AD	1.145	5.2441	7.02727
Error	AE	0.555	1.2321	1.65106
Error	AF	0.86	2.9584	3.96436
Error	AH	0.0425	0.007225	0.00968175
Error	BC	0.6275	1.57502	2.11059
Model	BG	-1.61	10.3684	13.894
	Lenth's ME	2.27978		
	Lenth's SME	4.62829		

The analysis of the data with the square root transformation identifies only *D*, *F*, the *BG* interaction as being significant. The original analysis identified factor *A* and several two factor interactions as being significant.



Design Expert Output

Response: sqrt

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	57.42	5	11.48	6.67	0.0056	significant
B	*0.56*	*1*	*0.56*	*0.32*	*0.5826*	
D	*18.71*	*1*	*18.71*	*10.87*	*0.0081*	
F	*27.20*	*1*	*27.20*	*15.81*	*0.0026*	
G	*0.59*	*1*	*0.59*	*0.34*	*0.5702*	
BG	*10.37*	*1*	*10.37*	*6.03*	*0.0340*	
Residual	17.21	10	1.72			
Cor Total	74.62	15				

The Model F-value of 6.67 implies the model is significant. There is only a 0.56% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.31	R-Squared	0.7694
Mean	2.32	Adj R-Squared	0.6541
C.V.	56.51	Pred R-Squared	0.4097
PRESS	44.05	Adeq Precision	8.422

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	2.32	1	0.33	1.59	3.05	
B-B	-0.19	1	0.33	-0.92	0.54	1.00
D-D	-1.08	1	0.33	-1.81	-0.35	1.00
F-F	-1.30	1	0.33	-2.03	-0.57	1.00
G-G	-0.19	1	0.33	-0.92	0.54	1.00
BG	-0.80	1	0.33	-1.54	-0.074	1.00

Final Equation in Terms of Coded Factors:

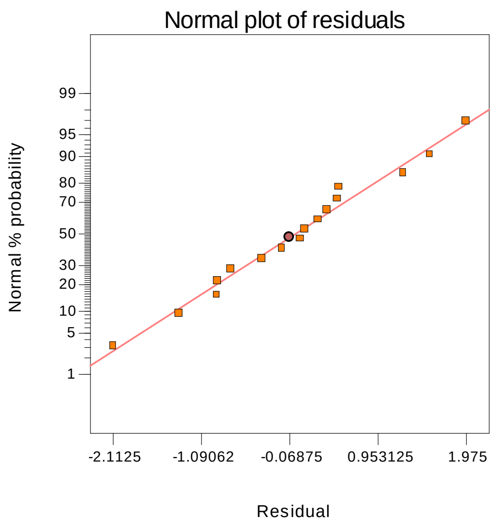
sqrt =
+2.32
-0.19 * B
-1.08 * D
-1.30 * F
-0.19 * G
-0.80 * B * G

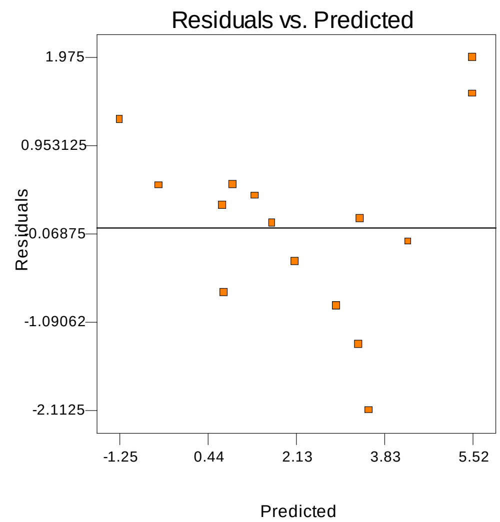
Final Equation in Terms of Actual Factors:

sqrt =

+2.32125	
-0.18625	* B
-1.08125	* D
-1.30375	* F
-0.19250	* G
-0.80500	* B * G

The residual plots are acceptable; although, there appears to be a slight "u" shape to the residuals versus predicted plot.





(f) There is a modification to the square root transformation proposed by Freeman and Tukey ("Transformations Related to the Angular and the Square Root," *Annals of Mathematical Statistics*, Vol. 21, 1950, pp. 607-611) that improves its performance. F&T's modification to the square root transformation is:

$$\frac{1}{2}\left[\sqrt{c}+\sqrt{c+1}\right]$$

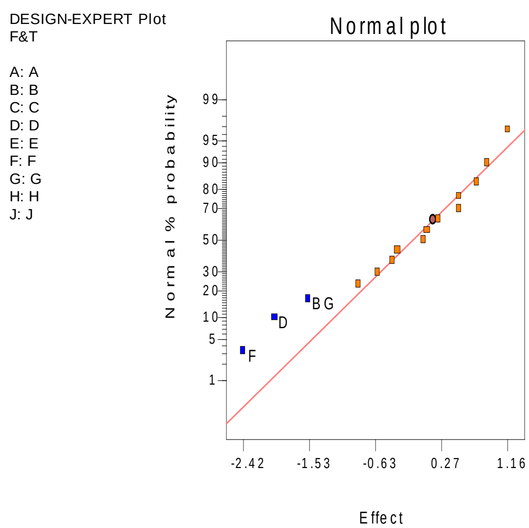
Rework parts (a) through (d) using this transformation and comment on the results. (For an interesting discussion and analysis of this experiment, refer to "Analysis of Factorial Experiments with Defects or Defectives as the Response," by S. Bisgaard and H.T. Fuller, *Quality Engineering*, Vol. 7, 1994-5, pp. 429-443.)

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Error	A	-0.86	2.9584	4.38512
Model	B	-0.325	0.4225	0.626255
Error	C	-0.605	1.4641	2.17018
Model	D	-1.995	15.9201	23.5977
Error	E	0.5025	1.01002	1.49712
Model	F	-2.425	23.5225	34.8664
Model	G	-0.4025	0.648025	0.960541
Error	H	0.225	0.2025	0.300158
Error	J	0.0275	0.003025	0.00448383
Error	AD	1.1625	5.40562	8.01254
Error	AE	0.505	1.0201	1.51205
Error	AF	0.8825	3.11523	4.61757
Error	AH	0.0725	0.021025	0.0311645
Error	BC	0.7525	2.26503	3.35735

Model	BG	-1.54	9.4864	14.0613
	Lenth's ME	2.14001		
	Lenth's SME	4.34453		

As with the square root transformation, factors *D*, *F*, and the *BG* interaction remain significant.



Design Expert Output

Response: F&T

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	50.00	5	10.00	5.73	0.0095	significant
B	*0.42*	*1*	*0.42*	*0.24*	*0.6334*	
D	*15.92*	*1*	*15.92*	*9.12*	*0.0129*	
F	*23.52*	*1*	*23.52*	*13.47*	*0.0043*	
G	*0.65*	*1*	*0.65*	*0.37*	*0.5560*	
BG	*9.49*	*1*	*9.49*	*5.43*	*0.0420*	
Residual	17.47	10	1.75			
Cor Total	67.46	15				

The Model F-value of 5.73 implies the model is significant. There is only a 0.95% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.32	R-Squared	0.7411
Mean	2.51	Adj R-Squared	0.6117
C.V.	52.63	Pred R-Squared	0.3373
PRESS	44.71	Adeq Precision	7.862

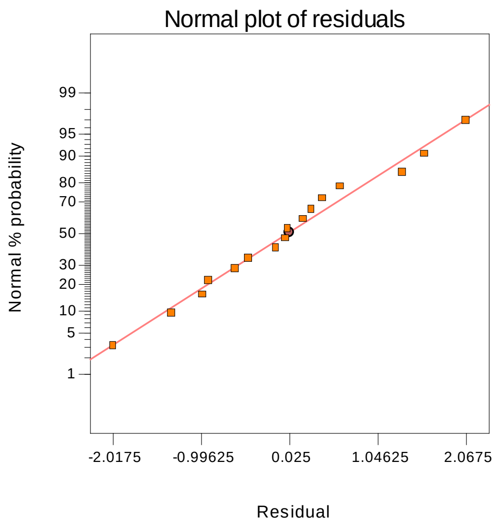
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	2.51	1	0.33	1.78	3.25	
B-B	-0.16	1	0.33	-0.90	0.57	1.00
D-D	-1.00	1	0.33	-1.73	-0.26	1.00
F-F	-1.21	1	0.33	-1.95	-0.48	1.00
G-G	-0.20	1	0.33	-0.94	0.53	1.00
BG	-0.77	1	0.33	-1.51	-0.034	1.00

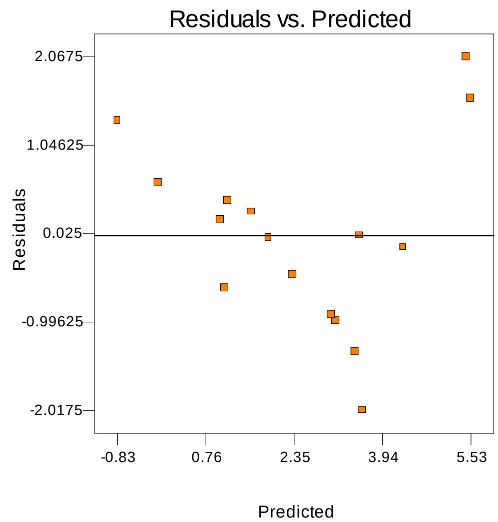
Final Equation in Terms of Coded Factors:

F&T =
+2.51
-0.16 * B
-1.00 * D

```
             -1.21     * F
             -0.20     * G
             -0.77     * B * G

     Final Equation in Terms of Actual Factors:

               F&T     =
          +2.51125
          -0.16250     * B
          -0.99750     * D
          -1.21250     * F
          -0.20125     * G
          -0.77000     * B * G
```

The following interaction plots appear as they did with the square root transformation; a slight “u” shape is observed in the residuals versus predicted plot.





8.52. An experiment is run in a semiconductor factory to investigate the effect of six factors on transistor gain. The design selected is the 2_{IV}^{6-2} shown in Table P8.15.

Table P8.15

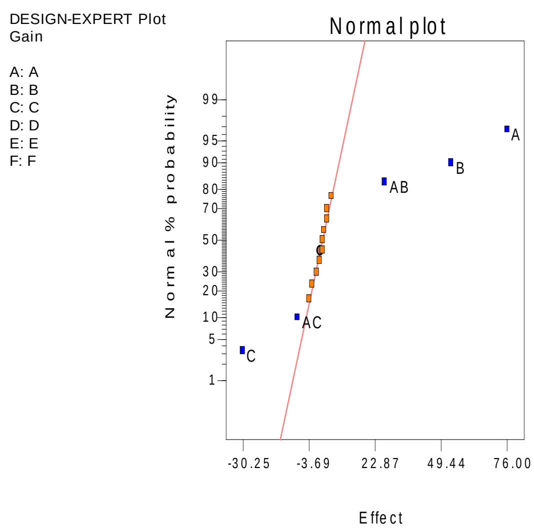
Standard Order	Run Order	*A*	*B*	*C*	*D*	*E*	*F*	Gain
1	2	-	-	-	-	-	-	1455
2	8	+	-	-	-	+	-	1511
3	5	-	+	-	-	+	+	1487
4	9	+	+	-	-	-	+	1596
5	3	-	-	+	-	+	+	1430
6	14	+	-	+	-	-	+	1481
7	11	-	+	+	-	-	-	1458
8	10	+	+	+	-	+	-	1549
9	15	-	-	-	+	-	+	1454
10	13	+	-	-	+	+	+	1517
11	1	-	+	-	+	+	-	1487
12	6	+	+	-	+	-	-	1596
13	12	-	-	+	+	+	-	1446
14	4	+	-	+	+	-	-	1473

15	7	-	+	+	+	-	+	1461
16	16	+	+	+	+	+	+	1563

(a) Use a normal plot of the effects to identify the significant factors.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	76	23104	55.2714
Model	B	53.75	11556.2	27.6459
Model	C	-30.25	3660.25	8.75637
Error	D	3.75	56.25	0.134566
Error	E	2	16	0.0382766
Error	F	1.75	12.25	0.0293055
Model	AB	26.75	2862.25	6.84732
Model	AC	-8.25	272.25	0.6513
Error	AD	-0.75	2.25	0.00538265
Error	AE	-3.5	49	0.117222
Error	AF	5.25	110.25	0.26375
Error	BD	0.5	1	0.00239229
Error	BF	2.5	25	0.0598072
Error	ABD	3.5	49	0.117222
Error	ABF	-2.5	25	0.0598072
	Lenth's ME	9.63968		
	Lenth's SME	19.57		



(b) Conduct appropriate statistical tests for the model identified in part (a).

Design Expert Output

Response: Gain
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	41455.00	5	8291.00	239.62	< 0.0001	significant
A	*23104.00*	*1*	*23104.00*	*667.75*	*< 0.0001*	
B	*11556.25*	*1*	*11556.25*	*334.00*	*< 0.0001*	
C	*3660.25*	*1*	*3660.25*	*105.79*	*< 0.0001*	
AB	*2862.25*	*1*	*2862.25*	*82.72*	*< 0.0001*	
AC	*272.25*	*1*	*272.25*	*7.87*	*0.0186*	
Residual	346.00	10	34.60			
Cor Total	41801.00	15				

The Model F-value of 239.62 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	5.88	R-Squared	0.9917
Mean	1497.75	Adj R-Squared	0.9876
C.V.	0.39	Pred R-Squared	0.9788
PRESS	885.76	Adeq Precision	44.419

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	1497.75	1	1.47	1494.47	1501.03	
A-A	38.00	1	1.47	34.72	41.28	1.00
B-B	26.87	1	1.47	23.60	30.15	1.00
C-C	-15.13	1	1.47	-18.40	-11.85	1.00
AB	13.38	1	1.47	10.10	16.65	1.00
AC	-4.12	1	1.47	-7.40	-0.85	1.00

Final Equation in Terms of Coded Factors:

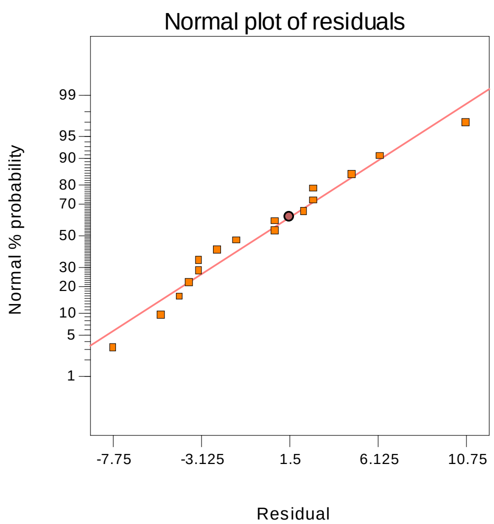
Gain	=
+1497.75	
+38.00	* A
+26.87	* B
-15.13	* C
+13.38	* A * B
-4.12	* A * C

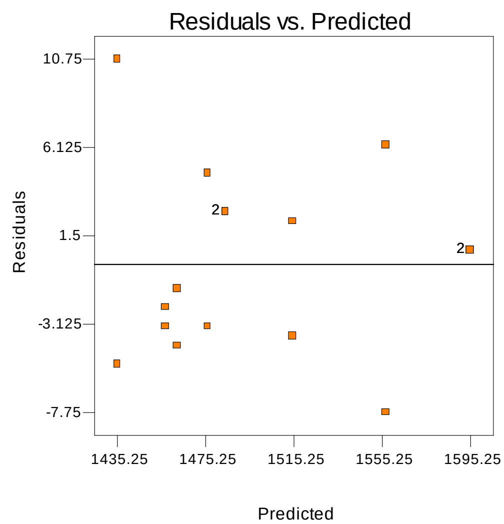
Final Equation in Terms of Actual Factors:

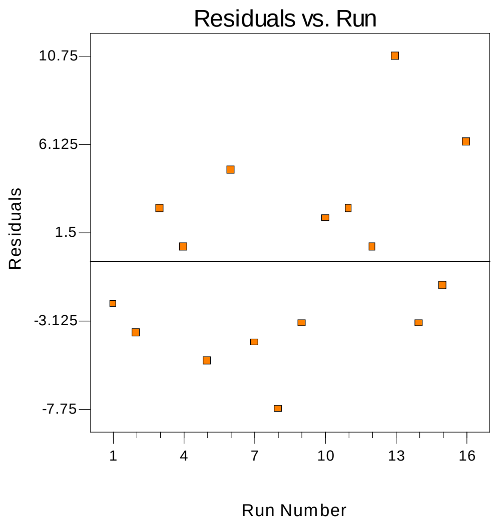
Gain	=
+1497.75000	
+38.00000	* A
+26.87500	* B
-15.12500	* C
+13.37500	* A * B
-4.12500	* A * C

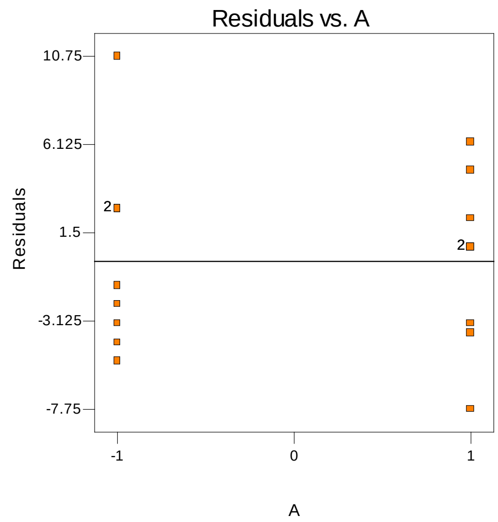
(c) Analyze the residuals and comment on your findings.

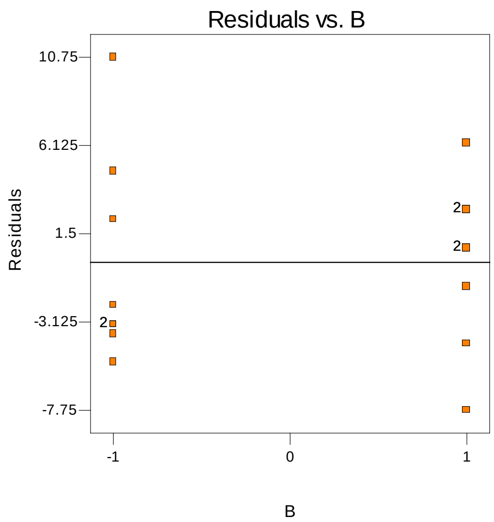
The residual plots are acceptable. The normality and equality of variance assumptions are verified. There does not appear to be any trends or interruptions in the residuals versus run order plot.

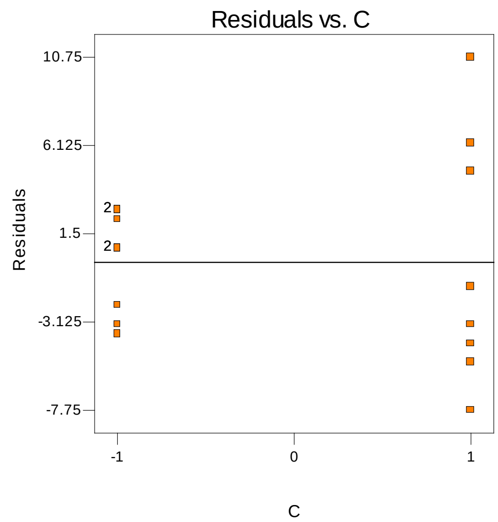


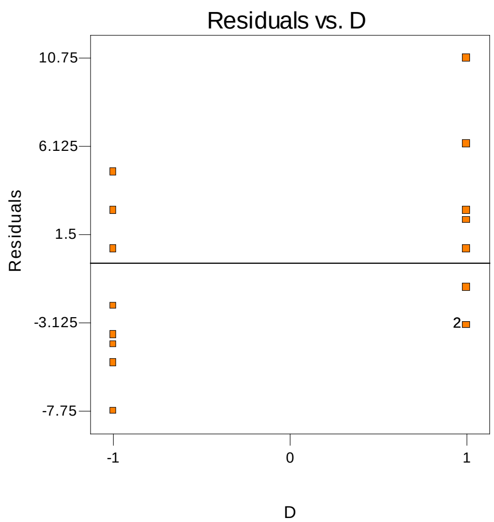


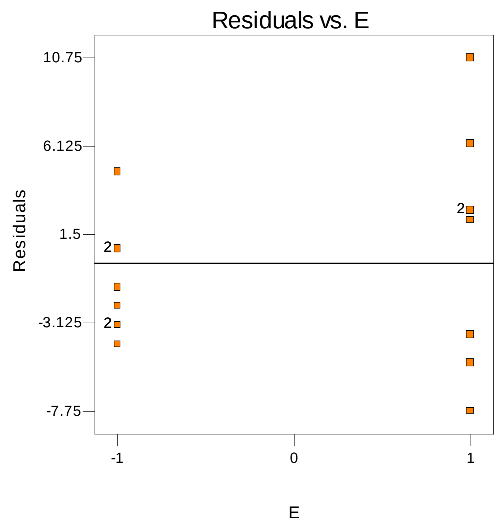

Residuals vs. Run
Residuals
10.75
6.125
1.5
-3.125
-7.75
1
4
7
10
13
16
Run Number

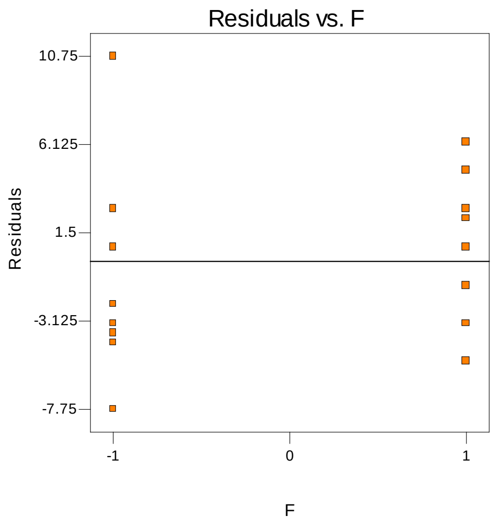

Residuals vs. A
Residuals
10.75
6.125
1.5
-3.125
-7.75
2
2
-1
0
1
A


Residuals vs. B
Residuals
10.75
6.125
1.5
-3.125
-7.75
2
2
2
-1
0
1
B


Residuals vs. C
Residuals
10.75
6.125
1.5
-3.125
-7.75
2
2
-1
0
1
C

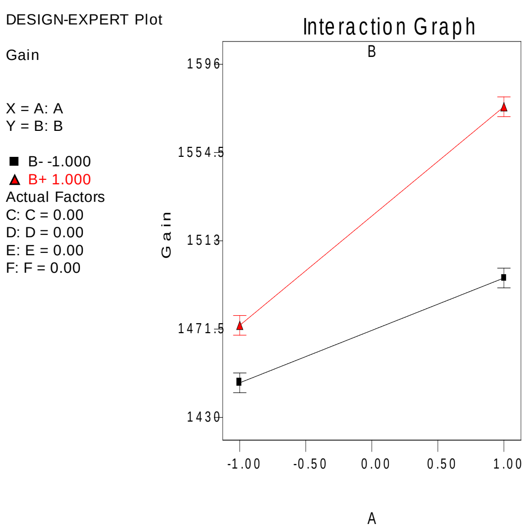

Residuals vs. D
Residuals
10.75
6.125
1.5
-3.125
-7.75
2
-1
0
1
D

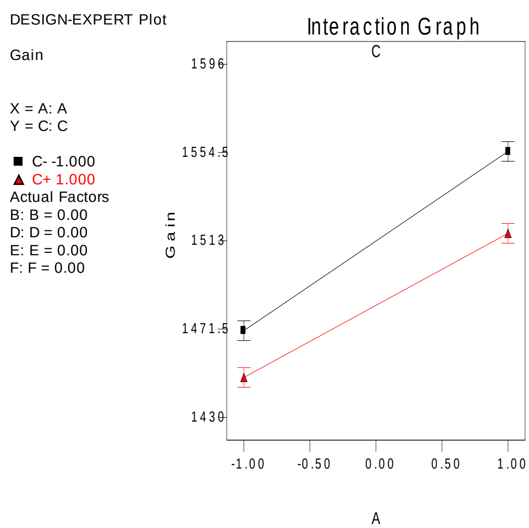

Residuals vs. E
Residuals
10.75
6.125
1.5
-3.125
-7.75
2
2
2
-1
0
1
E

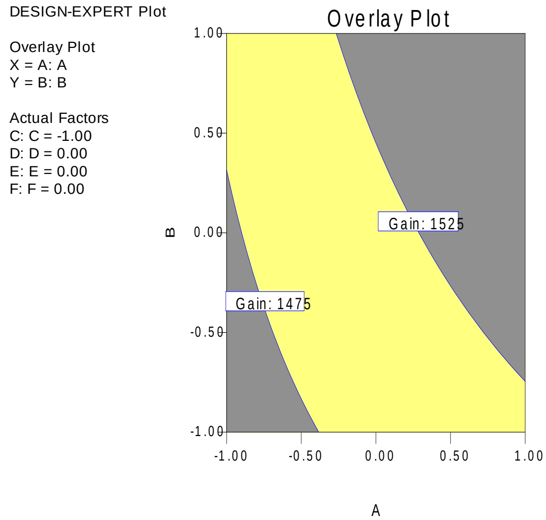


(d) Can you find a set of operating conditions that produce gain of 1500 ± 25 ?

Yes, see the graphs below.







8.53. Heat treating is often used to carbonize metal parts, such as gears. The thickness of the carbonized layer is a critical output variable from this process, and it is usually measured by performing a carbon analysis on the gear pitch (top of the gear tooth). Six factors were studied on a 2^{6-2}_{IV} design: A = furnace temperature, B = cycle time, C = carbon concentration, D = duration of the carbonizing cycle, E = carbon concentration of the diffuse cycle, and F = duration of the diffuse cycle. The experiment is shown in Table P8.16.

Table P8.16

Standard Order	Run Order	*A*	*B*	*C*	*D*	*E*	*F*	Pitch
1	5	-	-	-	-	-	-	74
2	7	+	-	-	-	+	-	190
3	8	-	+	-	-	+	+	133
4	2	+	+	-	-	-	+	127
5	10	-	-	+	-	+	+	115
6	12	+	-	+	-	-	+	101
7	16	-	+	+	-	-	-	54
8	1	+	+	+	-	+	-	144
9	6	-	-	-	+	-	+	121
10	9	+	-	-	+	+	+	188
11	14	-	+	-	+	+	-	135
12	13	+	+	-	+	-	-	170
13	11	-	-	+	+	+	-	126
14	3	+	-	+	+	-	-	175
15	15	-	+	+	+	-	+	126
16	4	+	+	+	+	+	+	193

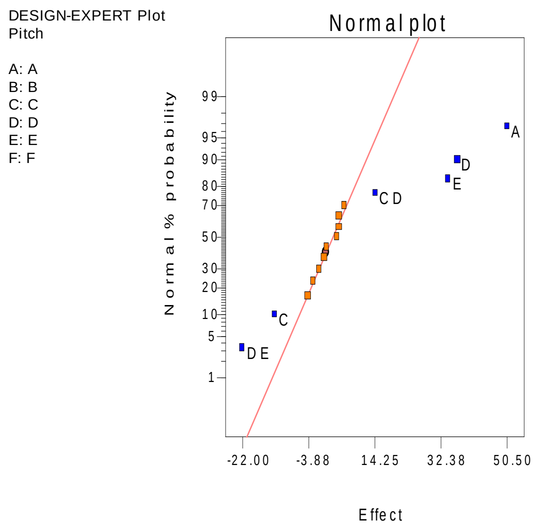
(a) Estimate the factor effects and plot them on a normal probability plot. Select a tentative model.

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Model	Intercept			
Model	A	50.5	10201	41.8777
Error	B	-1	4	0.016421
Model	C	-13	676	2.77515
Model	D	37	5476	22.4804

Model	E	34.5	4761	19.5451
Error	F	4.5	81	0.332526
Error	AB	-4	64	0.262737
Error	AC	-2.5	25	0.102631
Error	AD	4	64	0.262737
Error	AE	1	4	0.016421
Error	BD	4.5	81	0.332526
Model	CD	14.5	841	3.45252
Model	DE	-22	1936	7.94778
Error	ABD	0.5	1	0.00410526
Error	ABF	6	144	0.591157
	Lenth's ME	15.4235		
	Lenth's SME	31.3119		

Factors *A*, *C*, *D*, *E* and the two factor interactions *CD* and *DE* appear to be significant. The *CD* and *DE* interactions are aliased with BF and AF interactions respectively. Because factors B and F are not significant, *CD* and *DE* were included in the model. The model can be found in the Design Expert Output below.



(b) Perform appropriate statistical tests on the model.

Design Expert Output

Response: Pitch
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	23891.00	6	3981.83	76.57	< 0.0001	significant
A	*10201.00*	*1*	*10201.00*	*196.17*	*< 0.0001*	
C	*676.00*	*1*	*676.00*	*13.00*	*0.0057*	
D	*5476.00*	*1*	*5476.00*	*105.31*	*< 0.0001*	
E	*4761.00*	*1*	*4761.00*	*91.56*	*< 0.0001*	
CD	*841.00*	*1*	*841.00*	*16.17*	*0.0030*	
DE	*1936.00*	*1*	*1936.00*	*37.23*	*0.0002*	
Residual	468.00	9	52.00			
Cor Total	24359.00	15				

The Model F-value of 76.57 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	7.21	R-Squared	0.9808
Mean	135.75	Adj R-Squared	0.9680
C.V.	5.31	Pred R-Squared	0.9393

PRESS	1479.11		Adeq Precision	28.618	

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	135.75	1	1.80	131.67	139.83	
A-A	25.25	1	1.80	21.17	29.33	1.00
C-C	-6.50	1	1.80	-10.58	-2.42	1.00
D-D	18.50	1	1.80	14.42	22.58	1.00
E-E	17.25	1	1.80	13.17	21.33	1.00
CD	7.25	1	1.80	3.17	11.33	1.00
DE	-11.00	1	1.80	-15.08	-6.92	1.00

Final Equation in Terms of Coded Factors:

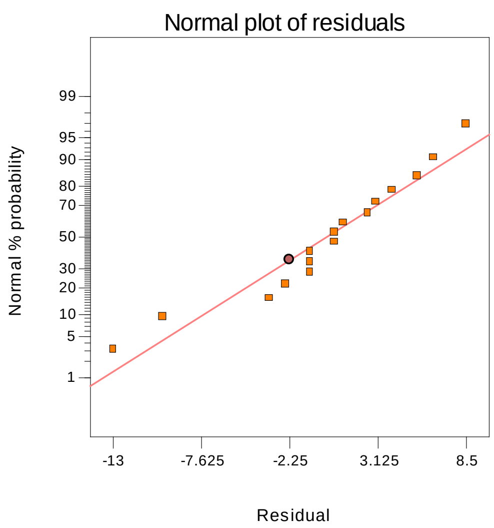
Pitch	=
+135.75	
+25.25	* A
-6.50	* C
+18.50	* D
+17.25	* E
+7.25	* C * D
-11.00	* D * E

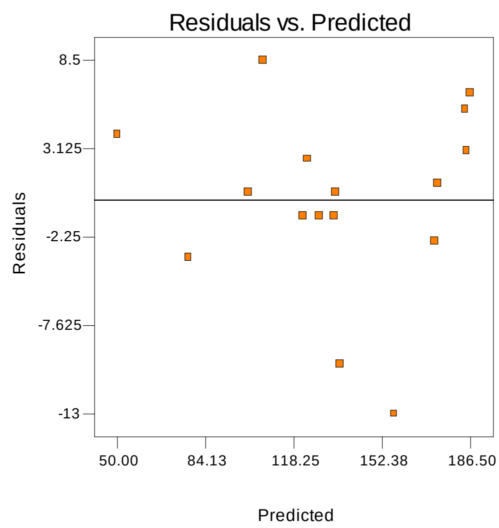
Final Equation in Terms of Actual Factors:

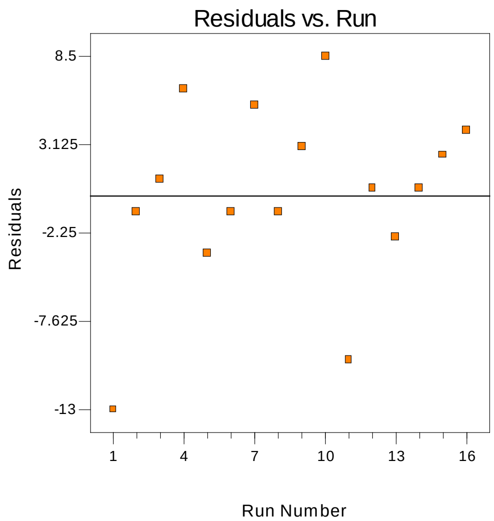
Pitch	=
+135.75000	
+25.25000	* A
-6.50000	* C
+18.50000	* D
+17.25000	* E
+7.25000	* C * D
-11.00000	* D * E

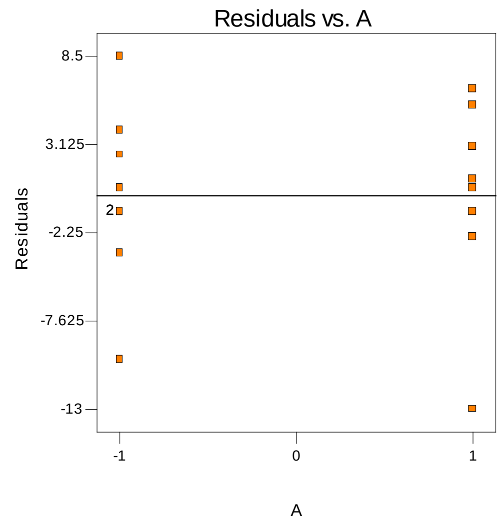
(c) Analyze the residuals and comment on model adequacy.

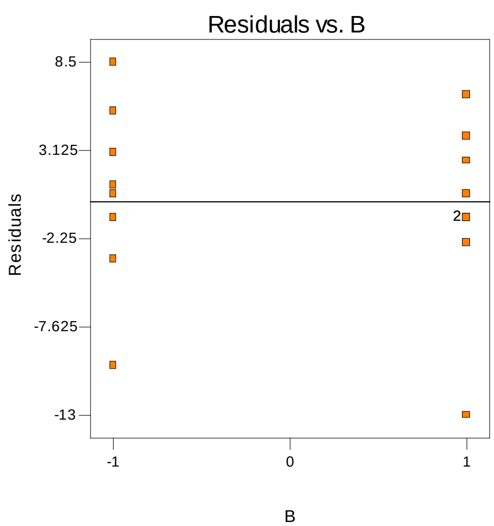
The residual plots are acceptable. The normality and equality of variance assumptions are verified. There does not appear to be any trends or interruptions in the residuals versus run order plot. The plots of the residuals versus factors *C* and *E* identify reduced variation at the lower level of both variables while the plot of residuals versus factor *F* identifies reduced variation at the upper level. Because *C* and *E* are significant factors in the model, this might not affect the decision on the optimum solution for the process. However, factor *F* is not included in the model and may be set at the upper level to reduce variation.

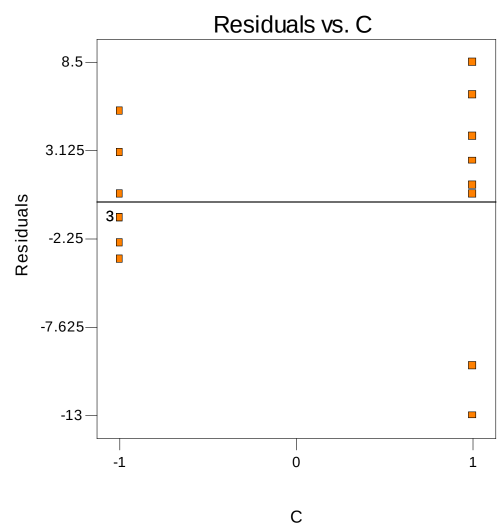


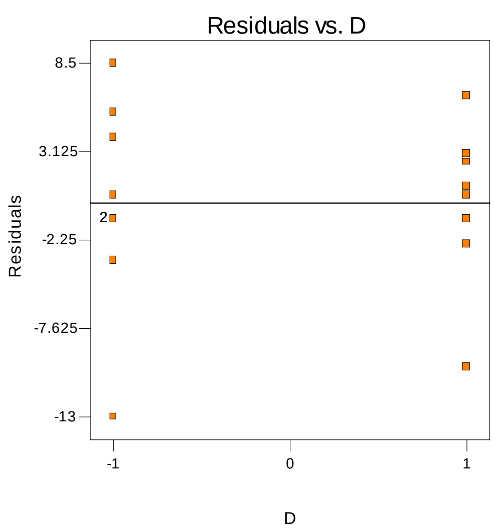


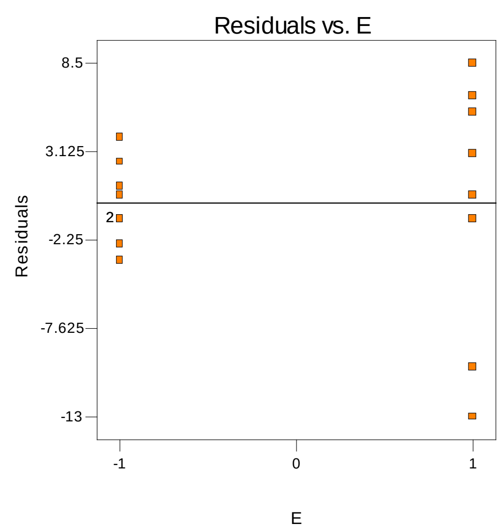

Residuals vs. Run
Residuals
8.5
3.125
-2.25
-7.625
-13
1 4 7 10 13 16
Run Number

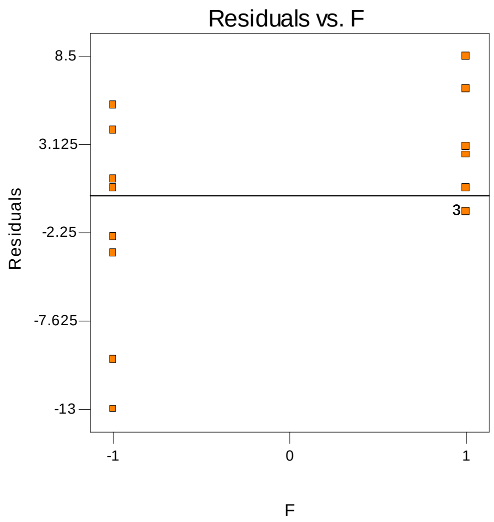

Residuals vs. A
Residuals
8.5
3.125
-2.25
-7.625
-13
2
-1 0 1
A


Residuals vs. B
Residuals
8.5
3.125
-2.25
-7.625
-13
2
-1 0 1
B


Residuals vs. C
Residuals
8.5
3.125
-2.25
-7.625
-13
3
-1 0 1
C

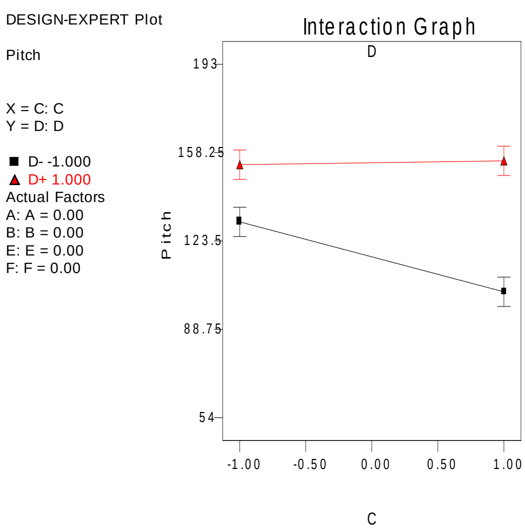

Residuals vs. D
Residuals
8.5
3.125
-2.25
-7.625
-13
2
-1 0 1
D

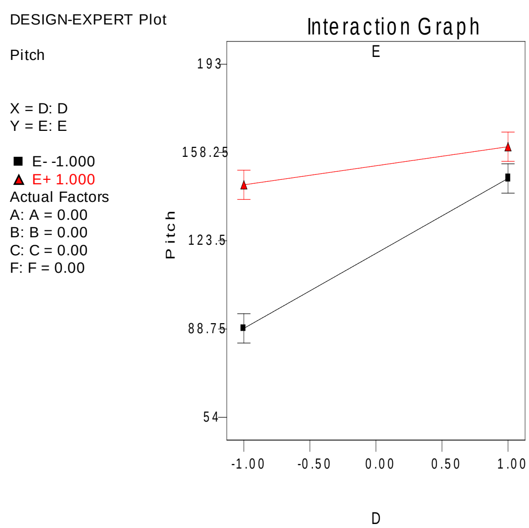

Residuals vs. E
Residuals
8.5
3.125
-2.25
-7.625
-13
2
-1 0 1
E

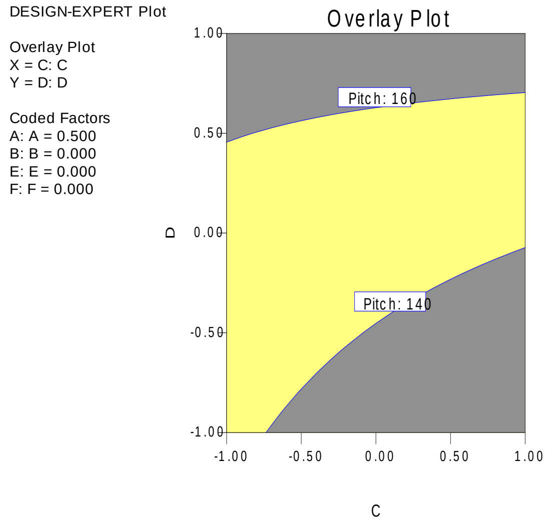


(d) Interpret the results of this experiment. Assume that a layer thickness of between 140 and 160 is desirable.

The graphs below identify a region that is acceptable between 140 and 160.







8.54. An article by L.B. Hare ("In the Soup: A Case Study to Identify Contributors to Filling Variability", Journal of Quality Technology, Vol. 20, pp. 36-43) describes a factorial experiment used to study the filling variability of dry soup mix packages. The factors are *A* = number of mixing ports through which the vegetable oil was added (1, 2), *B* = temperature surrounding the mixer (cooled, ambient), *C* = mixing time (60, 80 sec), *D* = batch weight (1500, 2000 lb), and *E* = number of days between mixing and packaging (1,7). Between 125 and 150 packages of soup were sampled over an eight hour period for each run in the design and the standard deviation of package weight was used as the response variable. The design and resulting data are shown in Table P8.17.

Table P8.17

Std Order	A - Mixer Ports	B - Temp	C - Time	D - Batch Weight	E - Delay	y - Std Dev
1	-1	-1	-1	-1	-1	1.13
2	1	-1	-1	-1	1	1.25
3	-1	1	-1	-1	1	0.97
4	1	1	-1	-1	-1	1.70
5	-1	-1	1	-1	1	1.47
6	1	-1	1	-1	-1	1.28
7	-1	1	1	-1	-1	1.18
8	1	1	1	-1	1	0.98
9	-1	-1	-1	1	1	0.78
10	1	-1	-1	1	-1	1.36
11	-1	1	-1	1	-1	1.85
12	1	1	-1	1	1	0.62
13	-1	-1	1	1	-1	1.09
14	1	-1	1	1	1	1.10
15	-1	1	1	1	1	0.76
16	1	1	1	1	-1	2.10

(a) What is the generator for this design?

The design generator is $I = -ABCDE$.

(b) What is the resolution of this design?

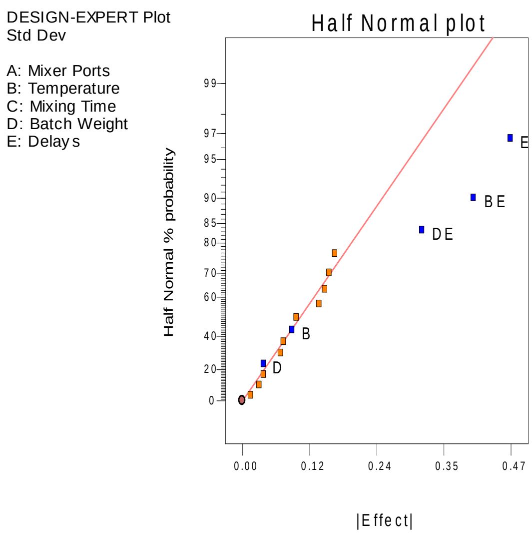
This design is Resolution *V*.

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

Design Expert Output

	Term	Effect	SumSqr	% Contribtn
Require	Intercept			
Error	A	0.145	0.0841	3.48388
Model	B	0.0875	0.030625	1.26865
Error	C	0.0375	0.005625	0.233018
Model	D	-0.0375	0.005625	0.233018
Model	E	-0.47	0.8836	36.6035
Error	AB	0.015	0.0009	0.0372829
Error	AC	0.095	0.0361	1.49546
Error	AD	0.03	0.0036	0.149132
Error	AE	-0.1525	0.093025	3.8536
Error	BC	-0.0675	0.018225	0.754979
Error	BD	0.1625	0.105625	4.37556
Model	BE	-0.405	0.6561	27.1792
Error	CD	0.0725	0.021025	0.87097
Error	CE	0.135	0.0729	3.01992
Model	DE	-0.315	0.3969	16.4418
	Lenth's ME	0.337389		
	Lenth's SME	0.684948		

Factor *E* and the two factor interactions *BE* and *DE* appear to be significant. Factors *B*, and *D*, are included to satisfy model hierarchy. The analysis of variance and model can be found in the Design Expert Output below.



Design Expert Output

Response: **Std Dev**
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	1.97	5	0.39	8.94	0.0019	significant
B	*0.031*	*1*	*0.031*	*0.69*	*0.4242*	
D	*5.625E-003*	*1*	*5.625E-003*	*0.13*	*0.7284*	
E	*0.88*	*1*	*0.88*	*20.03*	*0.0012*	
BE	*0.66*	*1*	*0.66*	*14.87*	*0.0032*	

DE	*0.40*	*1*	*0.40*	*9.00*	*0.0134*
Residual	0.44	10	0.044		
Cor Total	2.41	15			

The Model F-value of 8.94 implies the model is significant. There is only a 0.19% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.21	R-Squared	0.8173
Mean	1.23	Adj R-Squared	0.7259
C.V.	17.13	Pred R-Squared	0.5322
PRESS	1.13	Adeq Precision	9.252

Final Equation in Terms of Coded Factors:

Std Dev	=
+1.23	
+0.044	* B
-0.019	* D
-0.24	* E
-0.20	* B * E
-0.16	* D * E

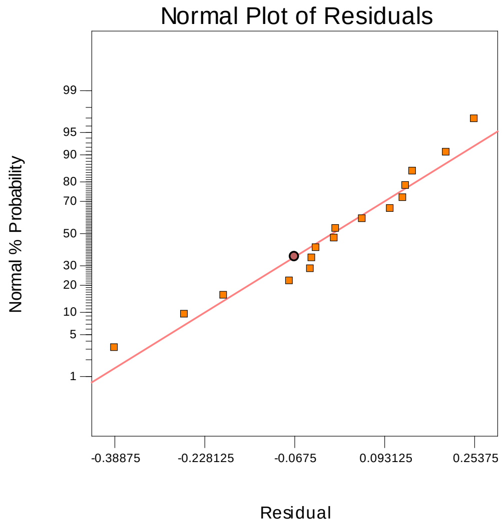
Final Equation in Terms of Actual Factors:

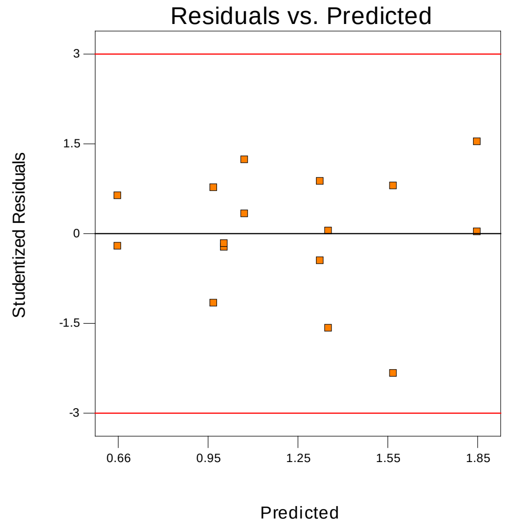
Temperature	Cool
Std Dev	=
-0.11292	
+7.65000E-004	* Batch Weight
+0.35667	* Delays
-2.10000E-004	* Batch Weight * Delays

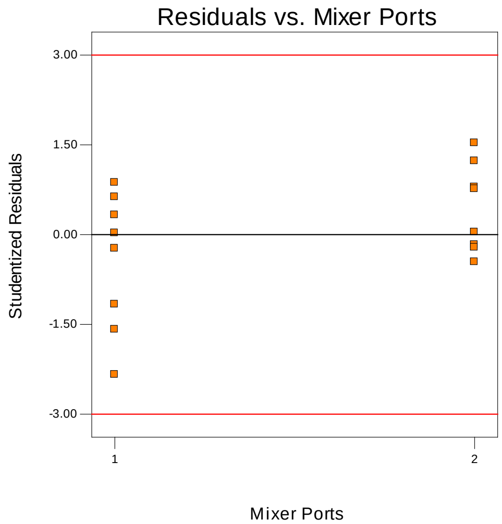
Temperature	Ambient
Std Dev	=
+0.51458	
+7.65000E-004	* Batch Weight
+0.22167	* Delays
-2.10000E-004	* Batch Weight * Delays

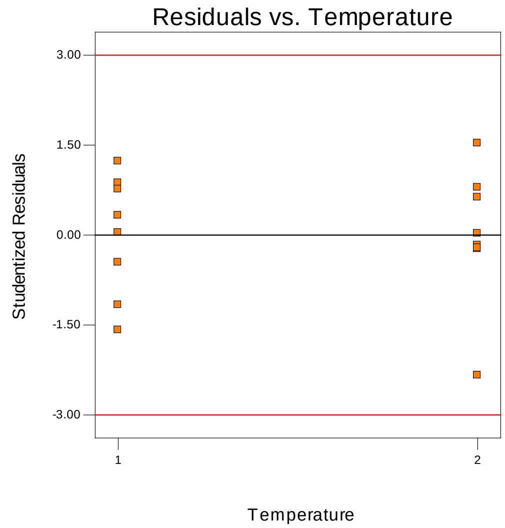
(d) Does a residual analysis indicate any problems with the underlying assumptions?

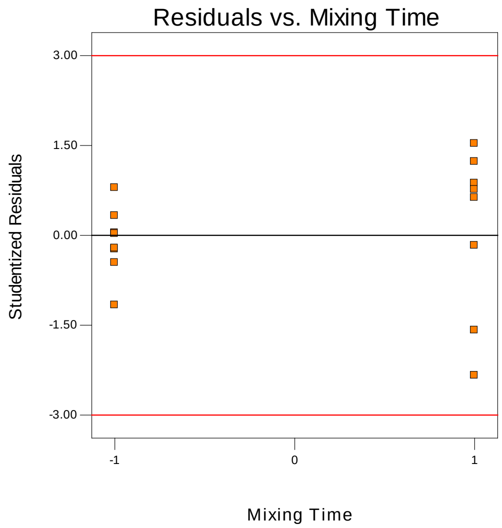
Often a transformation such as the natural log is required for the standard deviation response; however, the following residuals appear to be acceptable without the transformation.

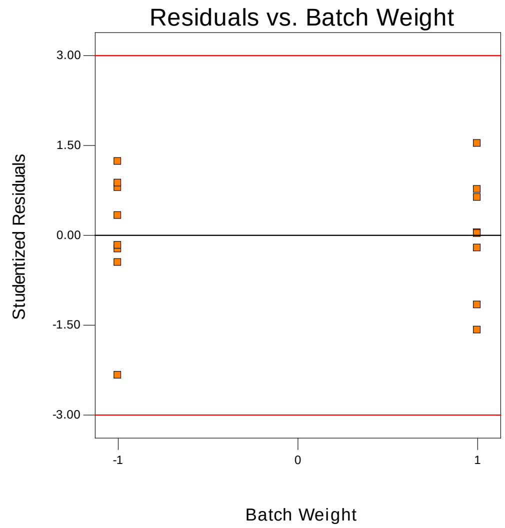

Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-0.38875
-0.228125
-0.0675
0.093125
0.25375
Residual

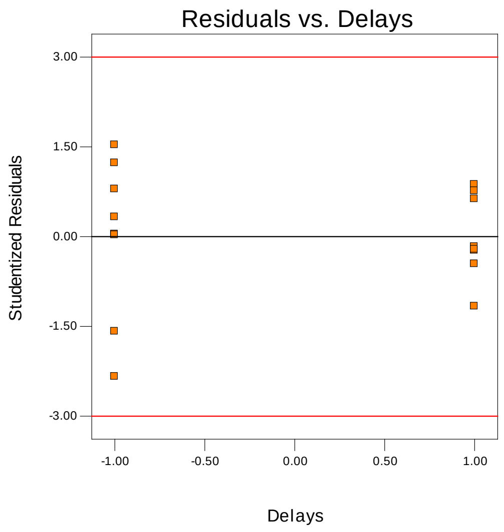

Residuals vs. Predicted
Studentized Residuals
3
1.5
0
-1.5
-3
0.66
0.95
1.25
1.55
1.85
Predicted


Residuals vs. Mixer Ports
Studentized Residuals
3.00
1.50
0.00
-1.50
-3.00
1
2
Mixer Ports


Residuals vs. Temperature
Studentized Residuals
3.00
1.50
0.00
-1.50
-3.00
1
2
Temperature

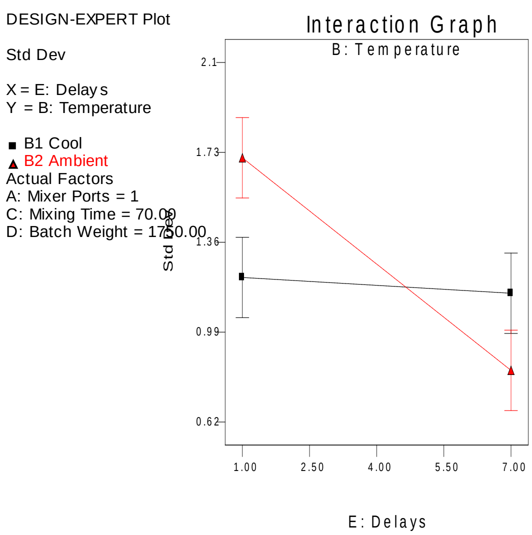

Residuals vs. Mixing Time
Studentized Residuals
3.00
1.50
0.00
-1.50
-3.00
-1
0
1
Mixing Time

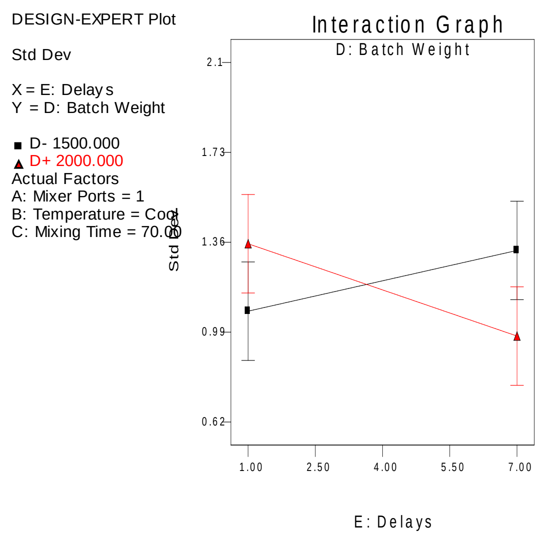

Residuals vs. Batch Weight
Studentized Residuals
3.00
1.50
0.00
-1.50
-3.00
-1
0
1
Batch Weight



(e) Draw conclusions about this filling process.

From the interaction plots below, the lowest standard deviation can be achieved with the Temperature at ambient, Batch Weight at 2000 lbs, and a Delay of 7 days.





8.55. Consider the 2_{IV}^{6-2} design.

(a) Suppose that the design had been folded over by changing the signs in the column B instead of column A. What changes would have resulted in the effects that can be estimated from the combined design?

Minitab Output – Fold Over on *A*

Fractional Factorial Design

```
Factors:   6   Base Design:        6, 16   Resolution:   IV
Runs:     32   Replicates:             1   Fraction:    1/2
Blocks:    1   Center pts (total):     0
```

```
Design Generators (before folding): E = ABC, F = BCD


Folded on Factors: A


Alias Structure

I + BCDF

A + ABCDF
B + CDF
C + BDF
D + BCF
E + BCDEF
F + BCD
AB + ACDF
AC + ABDF
AD + ABCF
AE + ABCDEF
AF + ABCD
BC + DF
BD + CF
BE + CDEF
BF + CD
CE + BDEF
DE + BCEF
EF + BCDE
ABC + ADF
ABD + ACF
ABE + ACDEF
ABF + ACD
ACE + ABDEF
ADE + ABCEF
AEF + ABCDE
BCE + DEF
BDE + CEF
BEF + CDE
ABCE + ADEF
ABDE + ACEF
ABEF + ACDE
```

Minitab Output – Fold Over on *B*

Fractional Factorial Design

```
Factors:   6   Base Design:         6, 16   Resolution:   IV
Runs:     32   Replicates:              1   Fraction:    1/2
Blocks:    1   Center pts (total):      0


Design Generators (before folding): E = ABC, F = BCD


Folded on Factors: B


Alias Structure

I + ADEF

A + DEF
B + ABDEF
C + ACDEF
D + AEF
E + ADF
F + ADE
AB + BDEF
AC + CDEF
AD + EF
AE + DF
```

```
AF + DE
BC + ABCDEF
BD + ABEF
BE + ABDF
BF + ABDE
CD + ACEF
CE + ACDF
CF + ACDE
ABC + BCDEF
ABD + BEF
ABE + BDF
ABF + BDE
ACD + CEF
ACE + CDF
ACF + CDE
BCD + ABCEF
BCE + ABCDF
BCF + ABCDE
ABCD + BCEF
ABCE + BCDF
ABCF + BCDE
```

Both combined designs are still resolution *IV*, there are some two-factor interactions aliased with other two-factor interactions. In the combined design folded on *A*, all two-factor interactions with *A* are now aliased with four-factor or higher interactions. The sparsity of effects principle would tell us that the higher order interactions are highly unlikely to occur. In the combined design folded on *B*, all two-factor interactions with *B* are now aliased with four-factor or higher interactions.

(b) Suppose that the design had been folded over by changing the sign in the column *E* instead of column *A*. What changes would have resulted in the effects that can be estimated from the combined design?

Minitab Output – Fold Over on *B*

Fractional Factorial Design

```
Factors:   6   Base Design:         6, 16   Resolution:   IV
Runs:     32   Replicates:              1   Fraction:    1/2
Blocks:    1   Center pts (total):      0

Design Generators (before folding): E = ABC, F = BCD

Folded on Factors: E

Alias Structure

I + BCDF

A + ABCDF
B + CDF
C + BDF
D + BCF
E + BCDEF
F + BCD
AB + ACDF
AC + ABDF
AD + ABCF
AE + ABCDEF
AF + ABCD
BC + DF
BD + CF
BE + CDEF
BF + CD
CE + BDEF
DE + BCEF
EF + BCDE
ABC + ADF
ABD + ACF
ABE + ACDEF
ABF + ACD
```

```
ACE + ABDEF
ADE + ABCEF
AEF + ABCDE
BCE + DEF
BDE + CEF
BEF + CDE
ABCE + ADEF
ABDE + ACEF
ABEF + ACDE
```

In the combined design folded on E, all two-factor interactions with E are now aliased with four-factor or higher interactions.

8.56. Consider the 2_{IV}^{7-3} design. Suppose that fold over of this design is run by changing the signs in column A. Determine the alias relationship in the combined design.

Minitab Output

```
Fractional Factorial Design

Factors:   7   Base Design:         7, 16   Resolution:   IV
Runs:     32   Replicates:              1   Fraction:    1/4
Blocks:    1   Center pts (total):      0

Design Generators (before folding): E = ABC, F = BCD, G = ACD

Folded on Factors: A

Alias Structure

I + BCDF + BDEG + CEFG

A + ABCDF + ABDEG + ACEFG
B + CDF + DEG + BCEFG
C + BDF + EFG + BCDEG
D + BCF + BEG + CDEFG
E + BDG + CFG + BCDEF
F + BCD + CEG + BDEFG
G + BDE + CEF + BCDFG
AB + ACDF + ADEG + ABCEFG
AC + ABDF + AEFG + ABCDEG
AD + ABCF + ABEG + ACDEFG
AE + ABDG + ACFG + ABCDEF
AF + ABCD + ACEG + ABDEFG
AG + ABDE + ACEF + ABCDFG
BC + DF + BEFG + CDEG
BD + CF + EG + BCDEFG
BE + DG + BCFG + CDEF
BF + CD + BCEG + DEFG
BG + DE + BCEF + CDFG
CE + FG + BCDG + BDEF
CG + EF + BCDE + BDFG
ABC + ADF + ABEFG + ACDEG
ABD + ACF + AEG + ABCDEFG
ABE + ADG + ABCFG + ACDEF
ABF + ACD + ABCEG + ADEFG
ABG + ADE + ABCEF + ACDFG
ACE + AFG + ABCDG + ABDEF
ACG + AEF + ABCDE + ABDFG
BCE + BFG + CDG + DEF
BCG + BEF + CDE + DFG
ABCE + ABFG + ACDG + ADEF
ABCG + ABEF + ACDE + ADFG
```

8.57. Reconsider the 2_{IV}^{7-3} design in Problem 8.56.

(a) Suppose that a fold over of this design is run by changing the signs in column *B*. Determine the alias relationship in the combined design.

Minitab Output

```
Fractional Factorial Design

Factors:   7   Base Design:         7, 16   Resolution:   IV
Runs:     32   Replicates:              1   Fraction:    1/4
Blocks:    1   Center pts (total):      0

Design Generators (before folding): E = ABC, F = BCD, G = ACD

Folded on Factors: B

Alias Structure

I + ACDG + ADEF + CEFG

A + CDG + DEF + ACEFG
B + ABCDG + ABDEF + BCEFG
C + ADG + EFG + ACDEF
D + ACG + AEF + CDEFG
E + ADF + CFG + ACDEG
F + ADE + CEG + ACDFG
G + ACD + CEF + ADEFG
AB + BCDG + BDEF + ABCEFG
AC + DG + AEFG + CDEF
AD + CG + EF + ACDEFG
AE + DF + ACFG + CDEG
AF + DE + ACEG + CDFG
AG + CD + ACEF + DEFG
BC + ABDG + BEFG + ABCDEF
BD + ABCG + ABEF + BCDEFG
BE + ABDF + BCFG + ABCDEG
BF + ABDE + BCEG + ABCDFG
BG + ABCD + BCEF + ABDEFG
CE + FG + ACDF + ADEG
CF + EG + ACDE + ADFG
ABC + BDG + ABEFG + BCDEF
ABD + BCG + BEF + ABCDEFG
ABE + BDF + ABCFG + BCDEG
ABF + BDE + ABCEG + BCDFG
ABG + BCD + ABCEF + BDEFG
ACE + AFG + CDF + DEG
ACF + AEG + CDE + DFG
BCE + BFG + ABCDF + ABDEG
BCF + BEG + ABCDE + ABDFG
ABCE + ABFG + BCDF + BDEG
ABCF + ABEG + BCDE + BDFG
```

(b) Compare the aliases from this combined design to those from the combined design from Problem 8.35. What differences resulted by changing the signs in a different column?

Both combined designs are still resolution *IV*, there are some two-factor interactions aliased with other two-factor interactions. In the combined design folded on *A*, all two-factor interactions with *A* are now aliased with four-factor or higher interactions. The sparsity of effects principle would tell us that the higher order interactions are highly unlikely to occur. In the combined design folded on *B*, all two-factor interactions with *B* are now aliased with four-factor or higher interactions.

8.58. Consider the 2_{IV}^{7-3} design.

(a) Suppose that a partial fold over of this design is run using column *A* (+ signs only). Determine the alias relationship in the combined design.

By choosing a fold over design in *Design Expert*, sorting on column *A*, and deleting the rows with a minus sign for *A* in the second block, the alias structures are identified below.

Design Expert Output

```
  Factorial Effects Aliases
[Est Terms]  Aliased Terms
   [Intercept] = Intercept + ABCE + ABFG + ACDG + ADEF + BCDF + BDEG + CEFG
   [A] = A - ABCE - ABFG - ACDG - ADEF + ABCDF + ABDEG + ACEFG
   [B] = B + ACE + AFG + CDF + DEG + ABCDG + ABDEF + BCEFG
   [C] = C + ABE + ADG + BDF + EFG + ABCFG + ACDEF + BCDEG
   [D] = D + ACG + AEF + BCF + BEG + ABCDE + ABDFG + CDEFG
   [E] = E + ABC + ADF + BDG + CFG + ABEFG + ACDEG + BCDEF
   [F] = F + ABG + ADE + BCD + CEG + ABCEF + ACDFG + BDEFG
   [G] = G + ABF + ACD + BDE + CEF + ABCEG + ADEFG + BCDFG
   [AB] = AB - ACE - AFG + ACDF + ADEG - ABCDG - ABDEF + ABCEFG
   [AC] = AC - ABE - ADG + ABDF + AEFG - ABCFG - ACDEF + ABCDEG
   [AD] = AD - ACG - AEF + ABCF + ABEG - ABCDE - ABDFG + ACDEFG
   [AE] = AE - ABC - ADF + ABDG + ACFG - ABEFG - ACDEG + ABCDEF
   [AF] = AF - ABG - ADE + ABCD + ACEG - ABCEF - ACDFG + ABDEFG
   [AG] = AG - ABF - ACD + ABDE + ACEF - ABCEG - ADEFG + ABCDFG
   [BC] = BC + DF + ABC + ADF + BEFG + CDEG + ABEFG + ACDEG
   [BD] = BD + CF + EG + ABCG + ABEF + ACDE + ADFG + BCDEFG
   [BE] = BE + DG + ABE + ADG + BCFG + CDEF + ABCFG + ACDEF
   [BF] = BF + CD + ABF + ACD + BCEG + DEFG + ABCEG + ADEFG
   [BG] = BG + DE + ABG + ADE + BCEF + CDFG + ABCEF + ACDFG
   [CE] = CE + FG + ACE + AFG + BCDG + BDEF + ABCDG + ABDEF
   [CG] = CG + EF + ACG + AEF + BCDE + BDFG + ABCDE + ABDFG
   [ABD] = ABD + ACF + AEG - ABCG - ABEF - ACDE - ADFG + ABCDEFG
   [BCE] = BCE + BFG + CDG + DEF + ABCE + ABFG + ACDG + ADEF
   [BCG] = BCG + BEF + CDE + DFG + ABCG + ABEF + ACDE + ADFG

   Factorial Effects Defining Contrast
   I = BCDF = BDEG = CEFG
```

(b) Rework part (a) using the negative signs to define the partial fold over. Does it make any difference which signs are used to define the partial fold over?

Both partial fold over designs produce the same alias relationships as shown below.

Design Expert Output

```
  Factorial Effects Aliases
[Est Terms]  Aliased Terms
   [Intercept] = Intercept + ABCE + ABFG + ACDG + ADEF + BCDF + BDEG + CEFG
   [A] = A - ABCE - ABFG - ACDG - ADEF + ABCDF + ABDEG + ACEFG
   [B] = B + ACE + AFG + CDF + DEG + ABCDG + ABDEF + BCEFG
   [C] = C + ABE + ADG + BDF + EFG + ABCFG + ACDEF + BCDEG
   [D] = D + ACG + AEF + BCF + BEG + ABCDE + ABDFG + CDEFG
   [E] = E + ABC + ADF + BDG + CFG + ABEFG + ACDEG + BCDEF
   [F] = F + ABG + ADE + BCD + CEG + ABCEF + ACDFG + BDEFG
   [G] = G + ABF + ACD + BDE + CEF + ABCEG + ADEFG + BCDFG
   [AB] = AB - ACE - AFG + ACDF + ADEG - ABCDG - ABDEF + ABCEFG
   [AC] = AC - ABE - ADG + ABDF + AEFG - ABCFG - ACDEF + ABCDEG
   [AD] = AD - ACG - AEF + ABCF + ABEG - ABCDE - ABDFG + ACDEFG
   [AE] = AE - ABC - ADF + ABDG + ACFG - ABEFG - ACDEG + ABCDEF
   [AF] = AF - ABG - ADE + ABCD + ACEG - ABCEF - ACDFG + ABDEFG
   [AG] = AG - ABF - ACD + ABDE + ACEF - ABCEG - ADEFG + ABCDFG
   [BC] = BC + DF + ABC + ADF + BEFG + CDEG + ABEFG + ACDEG
   [BD] = BD + CF + EG + ABCG + ABEF + ACDE + ADFG + BCDEFG
   [BE] = BE + DG + ABE + ADG + BCFG + CDEF + ABCFG + ACDEF
   [BF] = BF + CD + ABF + ACD + BCEG + DEFG + ABCEG + ADEFG
```

```
[BG] = BG + DE + ABG + ADE + BCEF + CDFG + ABCEF + ACDFG
[CE] = CE + FG + ACE + AFG + BCDG + BDEF + ABCDG + ABDEF
[CG] = CG + EF + ACG + AEF + BCDE + BDFG + ABCDE + ABDFG
[ABD] = ABD + ACF + AEG - ABCG - ABEF - ACDE - ADFG + ABCDEFG
[BCE] = BCE + BFG + CDG + DEF + ABCE + ABFG + ACDG + ADEF
[BCG] = BCG + BEF + CDE + DFG + ABCG + ABEF + ACDE + ADFG

Factorial Effects Defining Contrast
I = BCDF = BDEG = CEFG
```

8.59. Consider a partial fold over for the 2_{IV}^{6-2} design. Suppose that the signs are reversed in column *A*, but the eight runs are retained are the runs that have positive signs in column *C*. Determine the alias relationship in the combined design.

Design Expert Output

```
Factorial Effects Aliases
[Est Terms]  Aliased Terms
[Intercept] = Intercept - ABE + BCDF - ACDEF
[Block 1] = Block 1 + ABE + ABCE + ADEF + ACDEF
[Block 2] = Block 2 - ABE - ABCE - ADEF - ACDEF
[A] = A + BCE + DEF + ABCDF
[B] = B + ACE + CDF + ABDEF
[C] = C + ABE + BDF + ACDEF
[D] = D + BCF - ABDE - ACEF
[E] = E + ABC + ADF + BCDEF
[F] = F + BCD - ABEF - ACDE
[AB] = AB + ABC + ADF + ACDF
[AC] = AC - BCE - DEF + ABDF
[AD] = AD - BDE - CEF + ABCF
[AE] = AE + ACE + ABDEF + ABCDEF
[AF] = AF - BEF - CDE + ABCD
[BC] = BC + DF - ACE - ABDEF
[BD] = BD + CF + ABEF + ACDE
[BE] = BE + BCE + DEF + CDEF
[BF] = BF + CD + ABDE + ACEF
[CE] = CE - ABC - ADF + BDEF
[DE] = DE + BEF + CDE + BCEF
[EF] = EF + BDE + CEF + BCDE
[ABD] = ABD + ACF + BEF + CDE
[ABF] = ABF + ACD + BDE + CEF
[ADE] = ADE + ABEF + ACDE + ABCEF
[AEF] = AEF + ABDE + ACEF + ABCDE

Factorial Effects Defining Contrast
I = BCDF
```

8.60. Consider a partial fold over for the 2_{III}^{7-4} design. Suppose that the partial fold over of this design is constructed using column *A* (+ signs only). Determine the alias relationship for the combined design.

Design Expert Output

```
Factorial Effects Aliases
[Est Terms]  Aliased Terms
[Intercept] = Intercept - BD - CE - FG - ABCF + ABCG + ABEF - ABEG + ACDF
     - ACDG - ADEF + ADEG + BCDE + BDFG + CEFG - BCDEFG
[Block 1] = Block 1 + BD + CE + FG + ABD + ACE + AFG + BCF + BEG + CDG + DEF
     + ABCF + ABEG + ACDG + ADEF + BCDEFG + ABCDEFG
[Block 2] = Block 2 - BD - CE - FG - ABD - ACE - AFG - BCF - BEG - CDG - DEF
     - ABCF - ABEG - ACDG - ADEF - BCDEFG - ABCDEFG
[A] = A + BD + CE + FG + BCG + BEF + CDF + DEG + ABCF + ABEG + ACDG + ADEF
     + ABCDE + ABDFG + ACEFG + BCDEFG
[B] = B + AD + CF + EG + ACG + AEF + CDE + DFG + ABCE + ABFG + BCDG + BDEF
     + ABCDF + ABDEG + BCEFG + ACDEFG
```

```
[C] = C + AE + BF + DG + ABG + ADF + BDE + EFG + ABCD + ACFG + BCEG + CDEF
     + ABCEF + ACDEG + BCDFG + ABDEFG
[D] = D + AD + CF + EG + ACF + AEG + BCE + BFG + ABCE + ABFG + BCDG + BDEF
     + ABCDG + ABDEF + CDEFG + ACDEFG
[E] = E + AE + BF + DG + ABF + ADG + BCD + CFG + ABCD + ACFG + BCEG + CDEF
     + ABCEG + ACDEF + BDEFG + ABDEFG
[F] = F + AG + BC + DE + ABE + ACD + BDG + CEG + ABDF + ACEF + BEFG + CDFG
     + ABCFG + ADEFG + BCDEF + ABCDEG
[G] = G + AG + BC + DE + ABC + ADE + BDF + CEF + ABDF + ACEF + BEFG + CDFG
      + ABEFG + ACDFG + BCDEG + ABCDEG
[AB] = AB - AD - CF + CG + EF - EG - ABCE - ABFG + ACDE + ADFG + BCDF - BCDG
        - BDEF + BDEG + ABCEFG - ACDEFG
[AC] = AC - AE - BF + BG + DF - DG - ABCD + ABDE - ACFG + AEFG + BCEF - BCEG
        - CDEF + CDEG + ABCDFG - ABDEFG
[AF] = AF - AG - BC + BE + CD - DE - ABDF + ABDG - ACEF + ACEG + BCFG - BEFG
        - CDFG + DEFG + ABCDEF - ABCDEG

Factorial Effects Defining Contrast
   I = ABCG = ABEF = ACDF = ADEG = BCDE = BDFG = CEFG
```

8.61. Consider a partial fold over for the 2_{III}^{5-2} design. Suppose that the partial fold over of this design is constructed using column *A* (+ signs only). Determine the alias relationship for the combined design.

Design Expert Output

```
 Factorial Effects Aliases
[Est Terms]  Aliased Terms
 [Intercept] = Intercept + ABCE + ABFG + ACDG + ADEF + BCDF + BDEG + CEFG
    [Intercept] = Intercept - BD - CE + BCDE
 [Block 1] = Block 1 + BD + CE + ABD + ACE
 [Block 2] = Block 2 - BD - CE - ABD - ACE
 [A] = A + BD + CE + ABCDE
 [B] = B + AD + CDE + ABCE
 [C] = C + AE + BDE + ABCD
 [D] = D + AD + BCE + ABCE
 [E] = E + AE + BCD + ABCD
 [AB] = AB - AD - ABCE + ACDE
 [AC] = AC - AE - ABCD + ABDE
 [BC] = BC + DE + ABE + ACD
 [BE] = BE + CD + ABE + ACD
 [ABC] = ABC - ABE - ACD + ADE

 Factorial Effects Defining Contrast
 I = BCDE
```

8.62. Reconsider the 2^{4-1} design in Example 8.1. The significant factors are *A*, *C*, *D*, *AC* + *BD*, and *AD* + *BC*. Find a partial fold over design that will allow the *AC*, *BD*, *AD*, and *BD* interactions to be estimated.

By constructing a partial fold over reversing the signs in column *A* and using column *A* (+ signs only), the *AC*, *BD*, *AD*, and *BD* interactions can be estimated as shown below. This could also be accomplished by reversing the signs of any one of the factors *A*, *B*, *C*, and *D*.

Design Expert Output

```
 Factorial Effects Aliases
[Est Terms]  Aliased Terms
    [Intercept] = Intercept - BCD
 [Block 1] = Block 1 + BCD + ABCD
 [Block 2] = Block 2 - BCD - ABCD
 [A] = A + BCD
 [B] = B + ACD
 [C] = C + ABD
 [D] = D + ABC
 [AB] = AB - ACD
 [AC] = AC - ABD
```

```
[AD] = AD - ABC
[BC] = BC + ABC
[BD] = BD + ABD
[CD] = CD + ACD
```

8.63. Construct a supersaturated design for k = 8 factors in N = 6 runs.

We have chosen the Hadamard matrix design approach using the following Plackett-Burman design for N = 12 runs and k = 11 factors. By sorting on the 11^{th} factor, L, deleting the rows with the positive levels on L, and removing the last 3 factors, the design is reduced to 6 runs with 8 factors. Either the positive or negative levels on L could have been chosen. Also, the JMP statistics software package has the capability to generate supersaturated designs.

Run	*I*	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	*L*
1	+	+	+	-	+	+	+	-	-	-	+	-
2	+	-	+	+	-	+	+	+	-	-	-	+
3	+	+	-	+	+	-	+	+	+	-	-	-
4	+	-	+	-	+	+	-	+	+	+	-	-
5	+	-	-	+	-	+	+	-	+	+	+	-
6	+	-	-	-	+	-	+	+	-	+	+	+
7	+	+	-	-	-	+	-	+	+	-	+	+
8	+	+	+	-	-	-	+	-	+	+	-	+
9	+	+	+	+	-	-	-	+	-	+	+	-
10	+	-	+	+	+	-	-	-	+	-	+	+
11	+	+	-	+	+	+	-	-	-	+	-	+
12	+	-	-	-	-	-	-	-	-	-	-	-

Original Run	Run	*I*	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	1	1	1	1	-1	1	1	1	-1	-1
3	2	1	1	-1	1	1	-1	1	1	1
4	3	1	-1	1	-1	1	1	-1	1	1
5	4	1	-1	-1	1	-1	1	1	-1	1
9	5	1	1	1	1	-1	-1	-1	1	-1
12	6	1	-1	-1	-1	-1	-1	-1	-1	-1

8.63. Consider the 2^{8-3} design in problem 8.37. Suppose that the alias chain involving the AB interaction was large. Recommend a partial fold-over design to resolve the ambiguity about this interaction.

The full defining relationship is: I + ABCF + ABDG + CDFG + ADEFH + ACEGH + BCDEH; and the alias structure for the AB interaction is AB + CF + DG. The following 16 runs were created by changing the sign of A in the original experiment and then choosing the runs where A is at the low (-) level.

A	B	C	D	E	F	G	H
-1	-1	-1	-1	-1	1	1	1
-1	1	-1	-1	-1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1
-1	1	1	-1	-1	1	-1	1
-1	-1	-1	1	-1	1	-1	-1
-1	1	-1	1	-1	-1	1	1
-1	-1	1	1	-1	-1	-1	1
-1	1	1	1	-1	1	1	-1
-1	-1	-1	-1	1	1	1	-1
-1	1	-1	-1	1	-1	-1	1
-1	-1	1	-1	1	-1	1	1

-1	1	1	-1	1	1	-1	-1
-1	-1	-1	1	1	1	-1	1
-1	1	-1	1	1	-1	1	-1
-1	-1	1	1	1	-1	-1	-1
-1	1	1	1	1	1	1	1

8.64. Construct a supersaturated design for k = 12 factors in N = 10 runs.

We have chosen the Hadamard matrix design approach using the following Plackett-Burman design for N = 20 runs and k = 19 factors. By sorting on the 19th factor, T, deleting the rows with the negative levels on T, and removing the last 7 factors, the design is reduced to 10 runs with 12 factors. Either the positive or negative levels on T could have been chosen. Also, the JMP statistics software package has the capability to generate supersaturated designs.

Run	*I*	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*	*R*	*S*	*T*
1	+	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-
2	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+
3	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+
4	+	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-
5	+	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-
6	+	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-	-
7	+	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+	-
8	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-	+
9	+	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+	-
10	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-	+
11	+	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+	-
12	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+	+
13	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+	+
14	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+	+
15	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-	+
16	+	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-	-
17	+	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+	-
18	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+	+
19	+	+	-	-	+	+	+	+	-	+	-	+	-	-	-	-	+	+	-	+
20	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Original Run	Run	*I*	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	*L*	*M*
2	1	+	-	+	+	-	-	+	+	+	+	-	+	-
3	2	+	+	-	+	+	-	-	+	+	+	+	-	+
8	3	+	-	-	-	-	+	+	-	+	+	-	-	+
10	4	+	-	+	-	-	-	-	+	+	-	+	+	-
12	5	+	-	+	-	+	-	-	-	-	+	+	-	+
13	6	+	+	-	+	-	+	-	-	-	-	+	+	-
14	7	+	+	+	-	+	-	+	-	-	-	-	+	+
15	8	+	+	+	+	-	+	-	+	-	-	-	-	+
18	9	+	-	-	+	+	+	+	-	+	-	+	-	-
19	10	+	+	-	-	+	+	+	+	-	+	-	+	-

8.66. How could an "optimal design" approach be used to augment a fractional factorial design to de-alias effects of potential interest?

Both Design Expert and JMP software packages have the capability to augment designs and de-alias effects of potential interest. Upon completing the data entry and analysis in Design Expert for the fractional factorial experiment, design augmentation can be selected with the options of central composite, fold-over, and fractional factorial. By selecting fractional factorial, and choosing the additional model term(s) to be de-aliased, the D-optimal algorithm will insert additional runs as a new block for the experiment.

Chapter 9
Three-Level and Mixed-Level Factorial and Fractional Factorial Design Solutions

9.1. The effects of developer strength (A) and developer time (B) on the density of photographic plate film are being studied. Three strengths and three times are used, and four replicates of a 3^2 factorial experiment are run. The data from this experiment follow. Analyze the data using the standard methods for factorial experiments.

	Development Time (minutes)					
Developer Strength	10		14		18	
1	0	2	1	3	2	5
	5	4	4	2	4	6
2	4	6	6	8	9	10
	7	5	7	7	8	5
3	7	10	10	10	12	10
	8	7	8	7	9	8

Design Expert Output

Response: Data
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	224.22	8	28.03	10.66	< 0.0001	significant
A	*198.22*	*2*	*99.11*	*37.69*	*< 0.0001*	
B	*22.72*	*2*	*11.36*	*4.32*	*0.0236*	
AB	*3.28*	*4*	*0.82*	*0.31*	*0.8677*	
Pure Error	71.00	27	2.63			
Cor Total	295.22	35				

The Model F-value of 10.66 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.

Strength and time are significant. The quadratic and interaction effects are not significant. By treating both A and B as numerical factors, the analysis can be performed as follows:

Design Expert Output

Response: Data
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	214.71	2	107.35	44.00	< 0.0001	significant
A	*192.67*	*1*	*192.67*	*78.97*	*< 0.0001*	
B	*22.04*	*1*	*22.04*	*9.03*	*0.0050*	
Residual	80.51	33	2.44			
Lack of Fit	*9.51*	*6*	*1.59*	*0.60*	*0.7255*	*not significant*
Pure Error	*71.00*	*27*	*2.63*			
Cor Total	295.22	35				

The Model F-value of 44.00 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

9.2. Compute the I and J components of the two-factor interaction in Problem 9.1.

	B		
A	11	10	17
	22	28	32
	32	35	39

$$AB \text{ Totals} = 77, 78, 71;\quad I(AB) = \frac{77^2 + 78^2 + 71^2}{12} - \frac{226^2}{36} = 2.39$$

$$AB^2 \text{ Totals} = 78, 74, 74;\quad J(AB) = \frac{78^2 + 74^2 + 74^2}{12} - \frac{226^2}{36} = 0.89$$

$$SS_{AB} = I(AB) + J(AB) = 3.28$$

9.3. An experiment was performed to study the effect of three different types of 32-ounce bottles (A) and three different shelf types (B) -- smooth permanent shelves, end-aisle displays with grilled shelves, and beverage coolers -- on the time it takes to stock ten 12-bottle cases on the shelves. Three workers (factor C) were employed in this experiment, and two replicates of a 3^3 factorial design were run. The observed time data are shown in the following table. Analyze the data and draw conclusions.

		Replicate I			Replicate II		
Worker	Bottle Type	Permanent	End Aisle	Cooler	Permanent	End Aisle	Cooler
1	Plastic	3.45	4.14	5.80	3.36	4.19	5.23
	28-mm glass	4.07	4.38	5.48	3.52	4.26	4.85
	38-mm glass	4.20	4.26	5.67	3.68	4.37	5.58
2	Plastic	4.80	5.22	6.21	4.40	4.70	5.88
	28-mm glass	4.52	5.15	6.25	4.44	4.65	6.20
	38-mm glass	4.96	5.17	6.03	4.39	4.75	6.38
3	Plastic	4.08	3.94	5.14	3.65	4.08	4.49
	28-mm glass	4.30	4.53	4.99	4.04	4.08	4.59
	38-mm glass	4.17	4.86	4.85	3.88	4.48	4.90

Design Expert Output

Response: Time

ANOVA for Selected Factorial Model

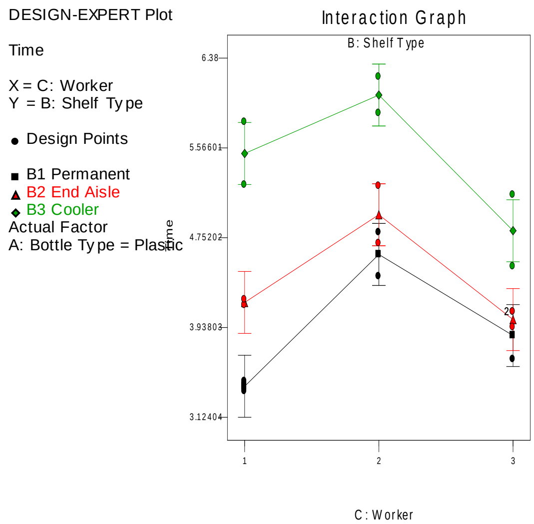
Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	28.28	26	1.09	14.50	< 0.0001	significant
A	*0.41*	*2*	*0.21*	*2.74*	*0.0828*	
B	*17.75*	*2*	*8.88*	*118.34*	*< 0.0001*	
C	*7.66*	*2*	*3.83*	*51.09*	*< 0.0001*	
AB	*0.12*	*4*	*0.029*	*0.39*	*0.8163*	
AC	*0.11*	*4*	*0.027*	*0.36*	*0.8319*	
BC	*1.68*	*4*	*0.42*	*5.60*	*0.0021*	
ABC	*0.55*	*8*	*0.069*	*0.92*	*0.5145*	
Pure Error	2.03	27	0.075			
Cor Total	30.31	53				

The Model F-value of 14.50 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, BC are significant model terms.

Factors *B* and *C*, shelf type and worker, and the *BC* interaction are significant. For the shortest time regardless of worker chose the permanent shelves. This can easily be seen in the interaction plot below.



9.4. A medical researcher is studying the effect of lidocaine on the enzyme level in the heart muscle of beagle dogs. Three different commercial brands of lidocaine (*A*), three dosage levels (*B*), and three dogs (*C*) are used in the experiment, and two replicates of a 3^3 factorial design are run. The observed enzyme levels follow. Analyze the data from this experiment.

		Replicate I			Replicate II		
		Dog			Dog		
Lidocaine Brand	Dosage Strength	1	2	3	1	2	3
1	1	96	84	85	84	85	86
	2	94	99	98	95	97	90
	3	101	106	98	105	104	103
2	1	85	84	86	80	82	84
	2	95	98	97	93	99	95
	3	108	114	109	110	102	100
3	1	84	83	81	83	80	79
	2	95	97	93	92	96	93
	3	105	100	106	102	111	108

Design Expert Output

Response: Enzyme Level

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	4346.26	26	167.16	13.10	< 0.0001	significant
A	*31.37*	*2*	*15.69*	*1.23*	*0.3083*	
B	*4040.04*	*2*	*2020.02*	*158.32*	*< 0.0001*	
C	*25.04*	*2*	*12.52*	*0.98*	*0.3879*	
AB	*112.52*	*4*	*28.13*	*2.20*	*0.0952*	
AC	*11.85*	*4*	*2.96*	*0.23*	*0.9178*	
BC	*56.52*	*4*	*14.13*	*1.11*	*0.3734*	
ABC	*68.93*	*8*	*8.62*	*0.68*	*0.7088*	
Pure Error	344.50	27	12.76			
Cor Total	4690.76	53				

```
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 B are significant model terms.
```

The factor *B*, dosage strength, is significant.

9.5. Compute the *I* and *J* components of the two-factor interactions for Example 9.1.

	A		
B	134	188	44
	-155	-348	-289
	176	127	288

I totals = 74, 75, 16 J totals = -128, 321, -28
I(*AB*) = 126.78 J(*AB*) = 6174.12
SS_{AB} = 6300.90

	A		
C	-190	-58	-211
	339	230	394
	6	-205	-140

I totals = -100, 342, -77 J totals = 25, 141, -1
I(*AC*) = 6878.78 J(*AC*) = 635.12
SS_{AC} = 7513.90

	B		
C	-93	-350	-16
	563	-133	533
	-104	-309	74

I totals = -152, 79, 238 J totals =-253, 287, 131
I(*BC*) = 4273.00 J(*BC*) = 8581.34
SS_{BC} = 12854.34

9.6. An experiment is run in a chemical process using a 3^2 factorial design. The design factors are temperature and pressure, and the response variable is yield. The data that result from this experiment are shown below.

	Pressure, psig		
Temperature, °C	100	120	140
80	47.58, 48.77	64.97, 69.22	80.92, 72.60
90	51.86, 82.43	88.47, 84.23	93.95, 88.54
100	71.18, 92.77	96.57, 88.72	76.58, 83.04

(a) Analyze the data from this experiment by conducting an analysis of variance. What conclusions can you draw?

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	3187.13	8	398.39	4.37	0.0205	significant
A	*1096.93*	*2*	*548.47*	*6.02*	*0.0219*	
B	*1503.56*	*2*	*751.78*	*8.25*	*0.0092*	
AB	*586.64*	*4*	*146.66*	*1.61*	*0.2536*	
Pure Error	819.98	9	91.11			
Cor Total	4007.10	17				

The Model F-value of 4.37 implies the model is significant. There is only a 2.05% 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.

Temperature and pressure are significant. Their interaction is not. An alternate analysis is performed below with *A* and *B* treated as numeric factors:

Design Expert Output

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

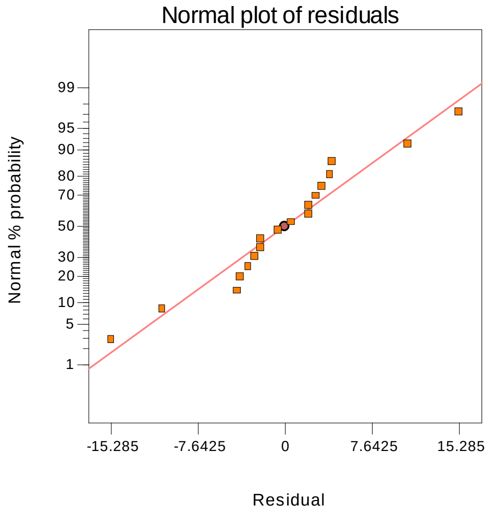
Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	3073.27	5	614.65	7.90	0.0017	significant
A	*850.76*	*1*	*850.76*	*10.93*	*0.0063*	
B	*1297.92*	*1*	*1297.92*	*16.68*	*0.0015*	
A2	*246.18*	*1*	*246.18*	*3.16*	*0.1006*	
B2	*205.64*	*1*	*205.64*	*2.64*	*0.1300*	
AB	*472.78*	*1*	*472.78*	*6.08*	*0.0298*	
Residual	933.83	12	77.82			
Lack of Fit	*113.86*	*3*	*37.95*	*0.42*	*0.7454*	*not significant*
Pure Error	*819.98*	*9*	*91.11*			
Cor Total	4007.10	17				

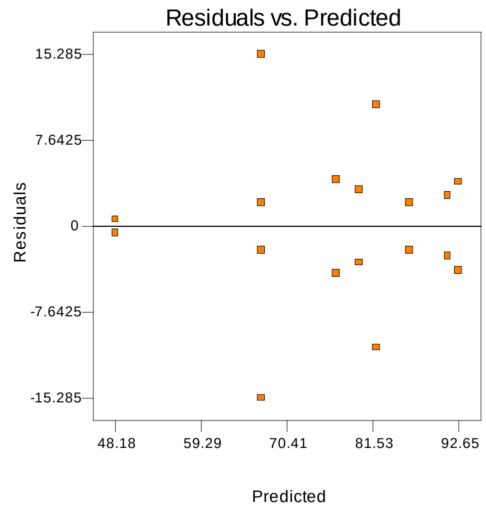
The Model F-value of 7.90 implies the model is significant. There is only a 0.17% chance that a "Model F-Value" this large could occur due to noise.

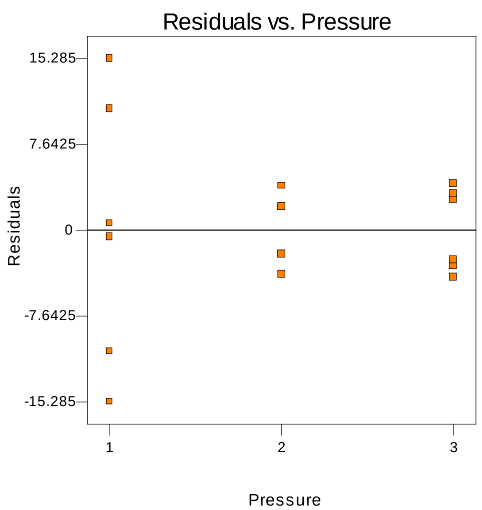
Values of "Prob > F" less than 0.0500 indicate model terms are significant. In this case A, B, AB are significant model terms.

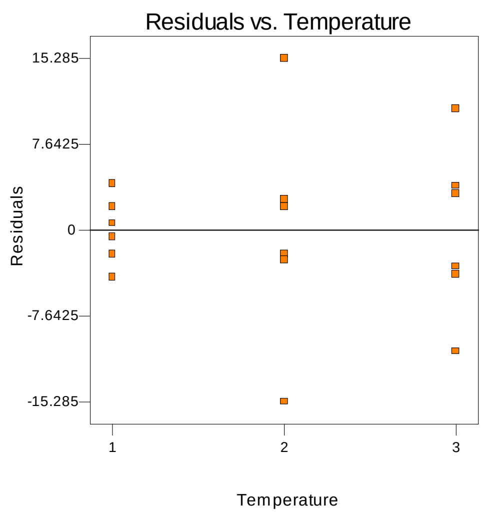
(b) Graphically analyze the residuals. Are there any concerns about underlying assumptions or model adequacy?

The following residual plots are based on the first analysis shown above with the *A* and *B* treated as categorical factors. The plot of residuals versus pressure shows a decreasing funnel shape indicating a non-constant variance.









(c) Verify that if we let the low, medium and high levels of both factors in this experiment take on the levels -1, 0, and +1, then a least squares fit to a second order model for yield is

$$\hat{y} = 86.81 + 10.4x_1 + 8.42x_2 - 7.17x_1^2 - 7.84x_2^2 - 7.69x_1x_2$$

The coefficients can be found in the following table of computer output.

Design Expert Output

Final Equation in Terms of Coded Factors:

```
Yield  =
    +86.81
     +8.42      * A
     +10.40     * B
     -7.84      * A^2
     -7.17      * B^2
     -7.69      * A * B
```

(d) Confirm that the model in part (c) can be written in terms of the natural variables temperature (T) and pressure (P) as

$$\hat{y} = -1335.63 + 18.56T + 8.59P - 0.072T^2 - 0.0196P^2 - 0.0384TP$$

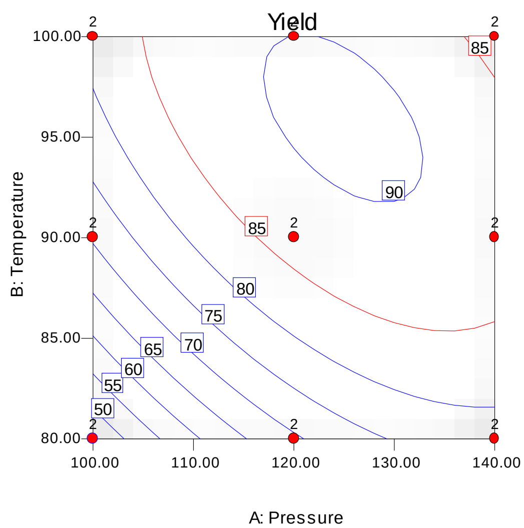
The coefficients can be found in the following table of computer output.

Design Expert Output

Final Equation in Terms of Actual Factors:

Yield	=
-1335.62500	
+8.58737	* Pressure
+18.55850	* Temperature
-0.019612	* Pressure^2
-0.071700	* Temperature^2
-0.038437	* Pressure * Temperature

(e) Construct a contour plot for yield as a function of pressure and temperature. Based on the examination of this plot, where would you recommend running the process?



Run the process in the oval region indicated by the yield of 90.

9.7.

(a) Confound a 3^3 design in three blocks using the ABC^2 component of the three-factor interaction. Compare your results with the design in Figure 9.7.

$$L = X_1 + X_2 + 2X_3$$

Block 1	Block 2	Block 3
000	100	200
112	212	012
210	010	110
120	220	020
022	122	222
202	002	102
221	021	121
101	201	001
011	111	211

The new design is a 180° rotation around the Factor *B* axis.

(b) Confound a 3^3 design in three blocks using the AB^2C component of the three-factor interaction. Compare your results with the design in Figure 9.7.

$$L = X_1 + 2X_2 + X_3$$

Block 1	Block 2	Block 3
000	210	112
022	202	120
011	221	101
212	100	010
220	122	002
201	111	021
110	012	200
102	020	222
121	001	211

The new design is a 180° rotation around the Factor *C* axis.

(c) Confound a 3^3 design three blocks using the ABC component of the three-factor interaction. Compare your results with the design in Figure 9.7.

$$L = X_1 + X_2 + X_3$$

Block 1	Block 2	Block 3
000	112	221
210	022	101
120	202	011
021	100	212
201	010	122
111	220	002
012	121	200
222	001	110
102	211	020

The new design is a 90° rotation around the Factor *C* axis along with switching layer 0 and layer 1 in the *C* axis.

(d) After looking at the designs in parts (a), (b), and (c) and Figure 9.7, what conclusions can you draw?

All four designs are relatively the same. The only differences are rotations and swapping of layers.

9.8. Confound a 3^4 design in three blocks using the AB^2CD component of the four-factor interaction.

The three blocks are shown below with L = $X_1 + 2X_2 + X_3 + X_4$

				Block 1				
0000	1100	0110	0101	2200	0220	0202	1210	1201
0211	1222	2212	2221	0122	2111	1121	1112	2010
2102	0021	2001	2120	1011	2022	0012	1002	1020

				Block 2				
1021	1110	1202	0001	0120	0212	1012	1101	1220
0200	0022	0111	2002	2121	2210	0010	0102	0221
1000	1122	1211	2112	2201	2020	2011	2100	2222

				Block 3				
2012	2101	2220	1022	1111	1200	2000	2121	2211
1221	1010	1102	0020	0112	0201	1001	1120	1212
2021	2110	2202	0100	0222	0011	0002	0121	0210

9.9. Consider the data from the first replicate of Problem 9.3. Assuming that all 27 observations could not be run on the same day, set up a design for conducting the experiment over three days with AB^2C confounded with blocks. Analyze the data.

	Block 1			Block 2			Block 3	
000	=	3.45	100	=	4.07	200	=	4.20
110	=	4.38	210	=	4.26	010	=	4.14
011	=	5.22	111	=	5.15	211	=	5.17
102	=	4.30	202	=	4.17	002	=	4.08
201	=	4.96	001	=	4.80	101	=	4.52
212	=	4.86	012	=	3.94	112	=	4.53
121	=	6.25	221	=	6.03	021	=	6.21
022	=	5.14	122	=	4.99	222	=	4.85
220	=	5.67	020	=	5.80	120	=	5.48
Totals	=	44.23			43.21			43.18

The analysis of variance below identifies factors B and C as significant.

Design Expert Output

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

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Block	0.079	2	0.040			
Model	13.57	18	0.75	10.49	0.0040	significant
A	*0.11*	*2*	*0.055*	*0.77*	*0.5052*	
B	*8.42*	*2*	*4.21*	*58.63*	*0.0001*	
C	*3.81*	*2*	*1.91*	*26.53*	*0.0010*	

AB	*0.30*	*4*	*0.075*	*1.04*	*0.4573*
AC	*0.11*	*4*	*0.028*	*0.39*	*0.8105*
BC	*0.81*	*4*	*0.20*	*2.83*	*0.1234*
Residual	0.43	6	0.072		
Cor Total	14.08	26			

The Model F-value of 10.49 implies the model is significant. There is only a 0.40% 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 are significant model terms.

9.10. Outline the analysis of variance table for the 3^4 design in nine blocks. Is this a practical design?

Source	DF
A	2
B	2
C	2
D	2
AB	4
AC	4
AD	4
BC	4
BD	4
CD	4
$ABC\ (AB^2C, ABC^2, AB^2C^2)$	6
$ABD\ (ABD, AB^2D, ABD^2)$	6
$ACD\ (ACD, ACD^2, AC^2D^2)$	6
$BCD\ (BCD, BC^2D, BCD^2)$	6
ABCD	16
Blocks ($ABC, AB^2C^2, AC^2D, BC^2D^2$)	8
Total	80

Any experiment with 81 runs is large. Instead of having three full levels of each factor, if two levels of each factor could be used, then the overall design would have 16 runs plus some center points. This two-level design could now probably be run in 2 or 4 blocks, with center points in each block. Additional curvature effects could be determined by augmenting the experiment with the axial points of a central composite design and additional enter points. The overall design would be less than 81 runs.

9.11. Consider the data in Problem 9.3. If *ABC* is confounded in replicate I and ABC^2 is confounded in replicate II, perform the analysis of variance.

$L_1 = X_1 + X_2 + X_3$ $\qquad\qquad$ $L_2 = X_1 + X_2 + 2X_2$

Block 1			Block 2			Block 3			Block 1			Block 2			Block 3		
000	=	3.45	001	=	4.80	002	=	4.08	000	=	3.36	100	=	3.52	200	=	3.68
111	=	5.15	112	=	4.53	110	=	4.38	101	=	4.44	201	=	4.39	001	=	4.40
222	=	4.85	220	=	5.67	221	=	6.03	011	=	4.70	111	=	4.65	211	=	4.75
120	=	5.48	121	=	6.25	122	=	4.99	221	=	6.38	021	=	5.88	121	=	6.20
102	=	4.30	100	=	4.07	101	=	4.52	202	=	3.88	002	=	3.65	102	=	4.04
210	=	4.26	211	=	5.17	212	=	4.86	022	=	4.49	122	=	4.59	222	=	4.90
201	=	4.96	202	=	4.17	200	=	4.20	120	=	4.85	220	=	5.58	020	=	5.23
012	=	3.94	010	=	4.14	011	=	5.22	210	=	4.37	010	=	4.19	110	=	4.26
021	=	6.21	022	=	5.14	020	=	5.80	112	=	4.08	212	=	4.48	012	=	4.08

The sums of squares for *A*, *B*, *C*, *AB*, *AC*, and *BC* are calculated as usual. The only sums of squares presenting difficulties with calculations are the four components of the *ABC* interaction (ABC, ABC^2, AB^2C, and AB^2C^2). ABC is computed using replicate I and ABC^2 is computed using replicate II. AB^2C and AB^2C^2 are computed using data from both replicates.

We will show how to calculate AB^2C and AB^2C^2 from both replicates. Form a two-way table of *A* x *B* at each level of *C*. Find the I(*AB*) and *J*(*AB*) totals for each third of the *A* x *B* table.

		A				
C	*B*	0	1	2	I	J
	0	6.81	7.59	7.88	26.70	27.55
0	1	8.33	8.64	8.63	27.25	27.17
	2	11.03	10.33	11.25	26.54	25.77
	0	9.20	8.96	9.35	31.41	31.24
1	1	9.92	9.80	9.92	30.97	31.29
	2	12.09	12.45	12.41	31.72	31.57
	0	7.73	8.34	8.05	26.09	26.29
2	1	8.02	8.61	9.34	27.31	26.11
	2	9.63	9.58	9.75	25.65	26.65

The *I* and J components for each third of the above table are used to form a new table of diagonal totals.

C		*I*(*AB*)			*J*(*AB*)	
0	26.70	27.25	26.54	27.55	27.17	25.77
1	31.41	30.97	31.72	31.24	31.29	31.57
2	26.09	27.31	25.65	26.29	26.11	26.65

I Totals:
85.06, 85.26, 83.32

I Totals:
85.49, 85.03, 83.12

J Totals:
85.73, 83.60, 84.31

J Totals:
83.35,8 5.06, 85.23

Now, $AB^2C^2 = I[C \text{ x } I(AB)] = \dfrac{(85.06)^2 + (85.26)^2 + (83.32)^2}{18} - \dfrac{(253.64)^2}{54} = 0.1265$

and, $AB^2C = J[C \text{ x } I(AB)] = \dfrac{(85.73)^2 + (83.60)^2 + (84.31)^2}{18} - \dfrac{(253.64)^2}{54} = 0.1307$

If it were necessary, we could find ABC^2 as $ABC^2 = I[C \text{ x } J(AB)]$ and ABC as $J[C \text{ x } J(AB)]$. However, these components must be computed using the data from the appropriate replicate.

The analysis of variance table:

Source	*SS*	*DF*	*MS*	F_0
Replicates	1.06696	1		
Blocks within Replicates	0.2038	4		
A	0.4104	2	0.2052	5.02
B	17.7514	2	8.8757	217.0
C	7.6631	2	3.8316	93.68
AB	0.1161	4	0.0290	<1
AC	0.1093	4	0.0273	<1
BC	1.6790	4	0.4198	10.26
ABC (rep I)	0.0452	2	0.0226	<1
ABC^2 (rep II)	0.1020	2	0.0510	1.25
AB^2C	0.1307	2	0.0754	1.60
AB^2C^2	0.1265	2	0.0633	1.55
Error	0.8998	22	0.0409	
Total	30.3069	53		

9.12. Consider the data from replicate I in Problem 9.3. Suppose that only a one-third fraction of this design with $I=ABC$ is run. Construct the design, determine the alias structure, and analyze the design.

The design is 000, 012, 021, 102, 201, 111, 120, 210, 222.

The alias structure is:

$$A = BC = AB^2C^2$$
$$B = AC = AB^2C$$
$$C = AB = ABC^2$$
$$AB^2 = AC^2 = BC^2$$

		C		
A	B	0	1	2
0	0	3.45		
	1			5.48
	2		4.26	
1	0			6.21
	1		5.15	
	2	4.96		
2	0		3.94	
	1	4.30		
	2			4.85

Source	SS	DF
A	2.25	2
B	0.30	2
C	2.81	2
AB^2	0.30	2
Total	5.66	8

9.13. From examining Figure 9.9, what type of design would remain if after completing the first 9 runs, one of the three factors could be dropped?

The remaining design is a full 3^2 factorial.

9.14. Construct a 3_{IV}^{4-1} design with $I=ABCD$. Write out the alias structure for this design.

The 27 runs for this design are as follows:

0000	1002	2001
0012	1011	2010
0021	1020	2022
0102	1101	2100
0111	1110	2112
0120	1122	2121
0201	1200	2202
0210	1212	2211
0222	1221	2220

$A = AB^2C^2D^2 = BCD$ $\quad B = AB^2CD = ACD$ $\quad C = ABC^2D = ABD$ $\quad D = ABCD^2 = ABC$
$AB = ABC^2D^2 = CD$ $\quad AB^2 = AC^2D^2 = BC^2D^2$ $\quad AC = AB^2CD^2 = BD$ $\quad AC^2 = AB^2D^2 = BC^2D$
$BC = AB^2C^2D = AD$ $\quad BC^2 = AB^2D = AC^2D$ $\quad BD^2 = AB^2C = ACD^2$ $\quad CD^2 = ABC^2 = ABD^2$
$AD^2 = AB^2C^2 = BCD^2$

9.15. Verify that the design in Problem 9.14 is a resolution IV design.

The design in Problem 9.14 is a Resolution IV design because no main effect is aliased with a component of a two-factor interaction, but some two-factor interaction components are aliased with each other.

9.16. Construct a 3^{5-2} design with $I=ABC$ and $I=CDE$. Write out the alias structure for this design. What is the resolution of this design?

The complete defining relation for this design is: $I = ABC = CDE = ABC^2DE = ABD^2E^2$
This is a resolution III design with $x_4 = 2x_1 + 2x_2 + x_3$ and $x_5 = x_1 + x_2 + 2x_4$ (mod 3).

00000	00112	00221
10022	10101	10210
20011	20120	20202
01022	01101	01210
11011	11120	11202
21000	21112	21221
02011	02120	02202
12000	12112	12221
22022	22101	22210

To find the alias of any effect, multiply the effect by I and I^2. For example, the alias of A is:

$A = AB^2C^2 = ACDE = AB^2CDE = AB^2DE = BC = AC^2D^2E^2 = BC^2DE = BD^2E^2$

9.17. Construct a 3^{9-6} design, and verify that it is a resolution III design.

Use the generators $I = AC^2D^2$, $I = AB^2C^2E$, $I = BC^2F^2$, $I = AB^2CG$, $I = ABCH^2$, and $I = ABJ^2$

000000000	021201102	102211001
022110012	212012020	001212210
011220021	100120211	211100110
221111221	122200220	020022222
210221200	010011111	222020101
202001212	201122002	200210122
112222112	002121120	121021010
101002121	111010202	110101022
120112100	220202011	012102201

To find the alias of any effect, multiply the effect by I and I^2. For example, the alias of C is:

$C = C(BC^2F^2) = BF^2$, At least one main effect is aliased with a component of a two-factor interaction.

9.18. Construct a 4×2^3 design confounded in two blocks of 16 observations each. Outline the analysis of variance for this design.

Design is a 4×2^3, with ABC at two levels, and Z at 4 levels. Represent Z with two pseudo-factors D and E as follows:

Factor Z	Pseudo- D	Factors E
Z_1	0	0 = (1)
Z_2	1	0 = d
Z_3	0	1 = e
Z_4	1	1 = de

The 4×2^3 is now a 2^5 in the factors A, B, C, D and E. Confound $ABCDE$ with blocks. We have given both the letter notation and the digital notation for the treatment combinations.

Block 1			Block 2		
(1)	=	0000	*a*	=	1000
ab	=	1100	*b*	=	0100
ac	=	1010	*c*	=	0010
bc	=	0110	*abc*	=	1110
abcd	=	1111	*bcd*	=	0111
abce	=	1112	*bce*	=	0112
cd	=	0011	*acd*	=	1011
ce	=	0012	*ace*	=	1012
de	=	0003	*ade*	=	1003
abde	=	1103	*bde*	=	0103
bcde	=	0113	*abcde*	=	1113
be	=	0102	*abd*	=	1101
ad	=	1001	*abe*	=	1102
ae	=	1002	*d*	=	0001
acde	=	1013	*e*	=	0002
bd	=	0101	*cde*	=	0013

Source	*DF*
A	1
B	1
C	1
Z (D+E+DE)	3
AB	1
AC	1
AZ (AD+AE+ADE)	3
BC	1
BZ (BD+BE+BDE)	3
CZ (CD+CE+CDE)	3
ABC	1
ABZ (ABD+ABE+ABDE)	3
ACZ (ACD+ACE+ACDE)	3
BCZ (BCD+BCE+BCDE)	3
ABCZ (ABCD+ABCE)	2
Blocks (or *ABCDE*)	1
Total	31

9.19. Outline the analysis of variance table for a $2^2 3^2$ factorial design. Discuss how this design may be confounded in blocks.

Suppose we have *n* replicates of a $2^2 3^2$ factorial design. *A* and *B* are at 2 levels, and *C* and *D* are at 3 levels.

Source	*DF*	Components for Confounding
A	1	*A*
B	1	*B*
C	2	*C*
D	2	*D*
AB	1	*AB*
AC	2	*AC*
AD	2	*AD*
BC	2	*BD*
BD	2	CD, CD^2
CD	4	*ABC*
ABC	2	*ABD*
ABD	2	ACD, ACD^2
ACD	4	BCD, BCD^2
BCD	4	$ABCD, ABCD^2$
ABCD	4	
Error	36(n-1)	
Total	36n-1	

Confounding in this series of designs is discussed extensively by Margolin (1967). The possibilities for a single replicate of the $2^2 3^2$ design are:

2 blocks of 18 observations	4 blocks of 9 observations	9 blocks of 4 observations
3 blocks of 12 observations	6 blocks of 6 observations	

For example, one component of the four-factor interaction, say $ABCD^2$, could be selected to confound the design in 3 blocks of 12 observations each, while to confound the design in 2 blocks of 18 observations each we would select the *AB* interaction. Cochran and Cox (1957) and Anderson and McLean (1974) discuss confounding in these designs.

9.20. Starting with a 16-run 2^4 design, show how two three-level factors can be incorporated in this experiment. How many two-level factors can be included if we want some information on two-factor interactions?

Use column *A* and *B* for one three-level factor and columns *C* and *D* for the other. Use the *AC* and *BD* columns for the two, two-level factors. The design will be of resolution V.

9.21. Starting with a 16-run 2^4 design, show how one three-level factor and three two-level factors can be accommodated and still allow the estimation of two-factor interactions.

Use columns *A* and *B* for the three-level factor, and columns *C* and *D* and *ABCD* for the three two-level factors. This design will be of resolution V.

9.22. In Problem 8.27, you met Harry Peterson-Nedry, a friend of the author who has a winery and vineyard in Newberg, Oregon. That problem described the application of two-level fractional factorial designs to their 1985 Pinor Noir product. In 1987, they wanted to conduct another Pinot Noir experiment. The variables for this experiment were

Variable	Levels
Clone of Pinot Noir	Wadenswil, Pommard
Berry Size	Small, Large
Fermentation temperature	80F, 85F, 90/80F, 90F
Whole Berry	None, 10%
Maceration Time	10 days, 21 days
Yeast Type	Assmanhau, Champagne
Oak Type	Troncais, Allier

Harry decided to use a 16-run two-level fractional factorial design, treating the four levels of fermentation temperature as two two-level variables. As in the Chapter 8 Problem, the rankings from a taste-test panel was the response variable. The design and the resulting average ranks are shown below:

Run	Clone	Size	Berry Temp.	Ferm. Berry	Whole Time	Macer Type	Yeast Type	Oak Rank	Average
1	-	-	-	-	-	-	-	-	4
2	+	-	-	-	-	+	+	+	10
3	-	+	-	-	+	-	+	+	6
4	+	+	-	-	+	+	-	-	9
5	-	-	+	-	+	+	+	-	11
6	+	-	+	-	+	-	-	+	1
7	-	+	+	-	-	+	-	+	15
8	+	+	+	-	-	-	+	-	5
9	-	-	-	+	+	+	-	+	12
10	+	-	-	+	+	-	+	-	2
11	-	+	-	+	-	+	+	-	16
12	+	+	-	+	-	-	-	+	3
13	-	-	+	+	-	-	+	+	8
14	+	-	+	+	-	+	-	-	14
15	-	+	+	+	+	-	-	-	7
16	+	+	+	+	+	+	+	+	13

(a) Describe the aliasing in this design.

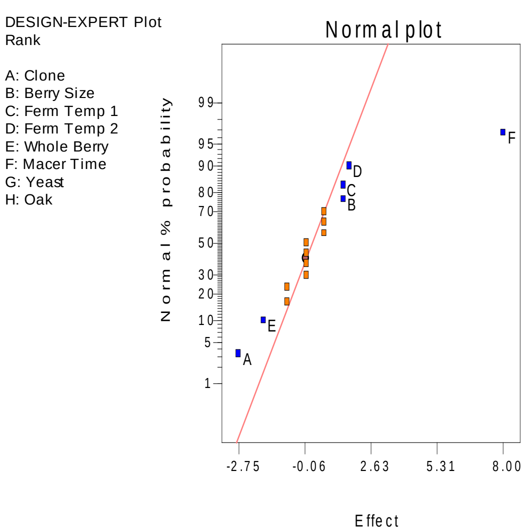
The design is a resolution IV design such that the main effects are aliased with three factor interactions.

Design Expert Output

Term	Aliases
Intercept	ABCG ABDH ABEF ACDF ACEH ADEG AFGH BCDE BCFH BDFG BEGH CDGH CEFG DEFH
A	BCG BDH BEF CDF CEH DEG FGH ABCDE
B	ACG ADH AEF CDE CFH DFG EGH
C	ABG ADF AEH BDE BFH DGH EFG
D	ABH ACF AEG BCE BFG CGH EFH
E	ABF ACH ADG BCD BGH CFG DFH
F	ABE ACD AGH BCH BDG CEG DEH
G	ABC ADE AFH BDF BEH CDH CEF
H	ABD ACE AFG BCF BEG CDG DEF
AB	CG DH EF ACDE ACFH ADFG AEGH BCDF BCEH BDEG BFGH
AC	BG DF EH ABDE ABFH ADGH AEFG BCDH BCEF CDEG CFGH
AD	BH CF EG ABCE ABFG ACGH AEFH BCDG BDEF CDEH DFGH
AE	BF CH DG ABCD ABGH ACFG ADFH BCEG BDEH CDEF EFGH
AF	BE CD GH ABCH ABDG ACEG ADEH BCFG BDFH CEFH DEFG
AG	BC DE FH ABDF ABEH ACDH ACEF BDGH BEFG CDFG CEGH
AH	BD CE FG ABCF ABEG ACDG ADEF BCGH BEFH CDFH DEGH

(b) Analyze the data and draw conclusions.

All of the main effects except Yeast and Oak are significant. The Macer Time is the most significant. None of the interactions were significant.



Design Expert Output

Response: Rank
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	328.75	6	54.79	43.83	< 0.0001	significant
A	*30.25*	*1*	*30.25*	*24.20*	*0.0008*	
B	*9.00*	*1*	*9.00*	*7.20*	*0.0251*	
C	*9.00*	*1*	*9.00*	*7.20*	*0.0251*	
D	*12.25*	*1*	*12.25*	*9.80*	*0.0121*	
E	*12.25*	*1*	*12.25*	*9.80*	*0.0121*	
F	*256.00*	*1*	*256.00*	*204.80*	*< 0.0001*	
Residual	11.25	9	1.25			
Cor Total	340.00	15				

The Model F-value of 43.83 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.12	R-Squared	0.9669	
Mean	8.50	Adj R-Squared	0.9449	
C.V.	13.15	Pred R-Squared	0.8954	
PRESS	35.56	Adeq Precision	19.270	

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	8.50	1	0.28	7.87	9.13	
A-Clone	-1.38	1	0.28	-2.01	-0.74	1.00
B-Berry Size	0.75	1	0.28	0.12	1.38	1.00
C-Ferm Temp 1	0.75	1	0.28	0.12	1.38	1.00
D-Ferm Temp 2	0.88	1	0.28	0.24	1.51	1.00
E-Whole Berry	-0.87	1	0.28	-1.51	-0.24	1.00
F-Macer Time	4.00	1	0.28	3.37	4.63	1.00

Final Equation in Terms of Coded Factors:

Rank	=
+8.50	
-1.38	* A
+0.75	* B
+0.75	* C
+0.88	* D
-0.87	* E
+4.00	* F

(c) What comparisons can you make between this experiment and the 1985 Pinot Noir experiment from Problem 8-27?

The experiment from Problem 8-27 indicates that yeast, barrel, whole cluster and the clone x yeast interactions were significant. This experiment indicates that maceration time, whole berry, clone and fermentation temperature are significant.

9.23. An article by W.D. Baten in the 1956 volume of *Industrial Quality Control* described an experiment to study the effect of three factors on the lengths of steel bars. Each bar was subjected to one of two heat treatment processes and was cut on one of four machines at one of three times during the day (8 am, 11 am, or 3 pm). The coded length data are shown below

Time of Day	Heat Treatment Process	Machine 1		Machine 2		Machine 3		Machine 4	
8am	1	6	9	7	9	1	2	6	6
		1	3	5	5	0	4	7	3
	2	4	6	6	5	-1	0	4	5
		0	1	3	4	0	1	5	4
11am	1	6	3	8	7	3	2	7	9
		1	-1	4	8	1	0	11	6
	2	3	1	6	4	2	0	9	4
		1	-2	1	3	-1	1	6	3
3pm	1	5	4	10	11	-1	2	10	5
		9	6	6	4	6	1	4	8
	2	6	0	8	7	0	-2	4	3
		3	7	10	0	4	-4	7	0

(a) Analyze the data from this experiment assuming that the four observations in each cell are replicates.

The Heat Treat Process effect (*B*) and Machine effect (*C*) are significant, while the Time of Day (*A*) is not significant. None of the interactions are significant.

Design Expert Output

Response: Length

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	590.33	23	25.67	4.13	< 0.0001	significant
A	*12.90*	*2*	*6.45*	*1.04*	*0.3596*	
B	*100.04*	*1*	*100.04*	*16.10*	*0.0001*	
C	*393.42*	*3*	*131.14*	*21.10*	*< 0.0001*	
AB	*1.65*	*2*	*0.82*	*0.13*	*0.8762*	
AC	*71.02*	*6*	*11.84*	*1.90*	*0.0917*	
BC	*1.54*	*3*	*0.51*	*0.083*	*0.9693*	
ABC	*9.77*	*6*	*1.63*	*0.26*	*0.9527*	
Pure Error	447.50	72	6.22			
Cor Total	1037.83	95				

The Model F-value of 4.13 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.49	R-Squared	0.5688
Mean	3.96	Adj R-Squared	0.4311
C.V.	62.98	Pred R-Squared	0.2334
PRESS	795.56	Adeq Precision	7.020

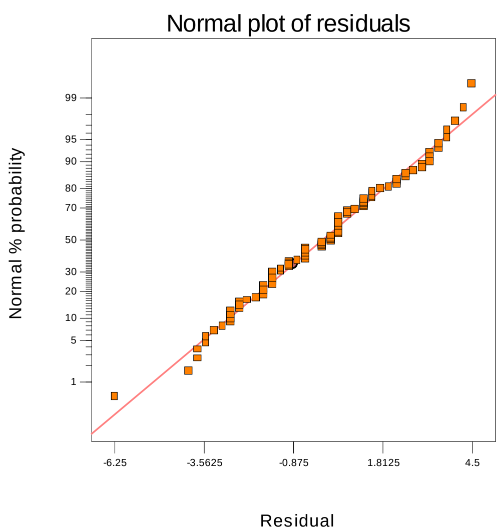
Term	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	3.96	1	0.25	3.45	4.47	
A[1]	-0.18	1	0.36	-0.89	0.54	
A[2]	-0.33	1	0.36	-1.05	0.38	
B-Process	-1.02	1	0.25	-1.53	-0.51	1.00
C[1]	-0.54	1	0.44	-1.42	0.34	
C[2]	1.92	1	0.44	1.04	2.80	
C[3]	-3.08	1	0.44	-3.96	-2.20	
A[1]B	0.18	1	0.36	-0.54	0.89	
A[2]B	-0.042	1	0.36	-0.76	0.68	
A[1]C[1]	0.51	1	0.62	-0.73	1.75	
A[2]C[1]	-1.58	1	0.62	-2.83	-0.34	
A[1]C[2]	-0.20	1	0.62	-1.44	1.04	
A[2]C[2]	-0.42	1	0.62	-1.66	0.83	
A[1]C[3]	0.18	1	0.62	-1.07	1.42	
A[2]C[3]	0.46	1	0.62	-0.78	1.70	
BC[1]	0.10	1	0.44	-0.77	0.98	
BC[2]	-0.10	1	0.44	-0.98	0.77	
BC[3]	0.15	1	0.44	-0.73	1.02	
A[1]BC[1]	-0.26	1	0.62	-1.50	0.98	
A[2]BC[1]	0.21	1	0.62	-1.03	1.45	
A[1]BC[2]	-0.052	1	0.62	-1.29	1.19	
A[2]BC[2]	-0.46	1	0.62	-1.70	0.78	
A[1]BC[3]	-0.18	1	0.62	-1.42	1.07	
A[2]BC[3]	0.42	1	0.62	-0.83	1.66	

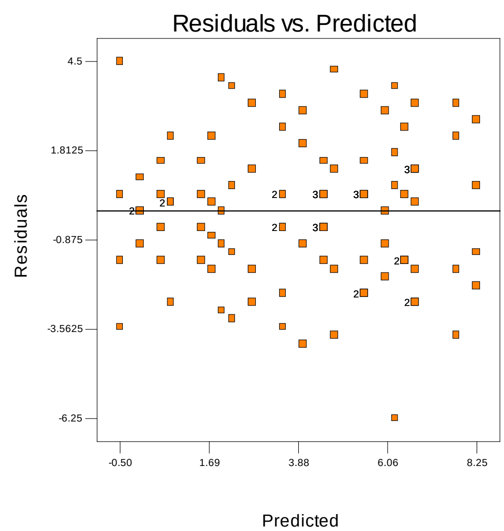
Final Equation in Terms of Coded Factors:

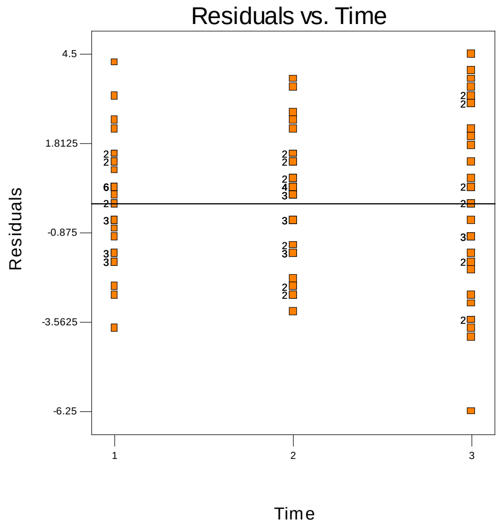
Length	=
+3.96	
-0.18	* A[1]
-0.33	* A[2]
-1.02	* B
-0.54	* C[1]
+1.92	* C[2]
-3.08	* C[3]
+0.18	* A[1]B
-0.042	* A[2]B
+0.51	* A[1]C[1]
-1.58	* A[2]C[1]
-0.20	* A[1]C[2]
-0.42	* A[2]C[2]
+0.18	* A[1]C[3]
+0.46	* A[2]C[3]
+0.10	* BC[1]
-0.10	* BC[2]
+0.15	* BC[3]
-0.26	* A[1]BC[1]
+0.21	* A[2]BC[1]
-0.052	* A[1]BC[2]
-0.46	* A[2]BC[2]
-0.18	* A[1]BC[3]
+0.42	* A[2]BC[3]

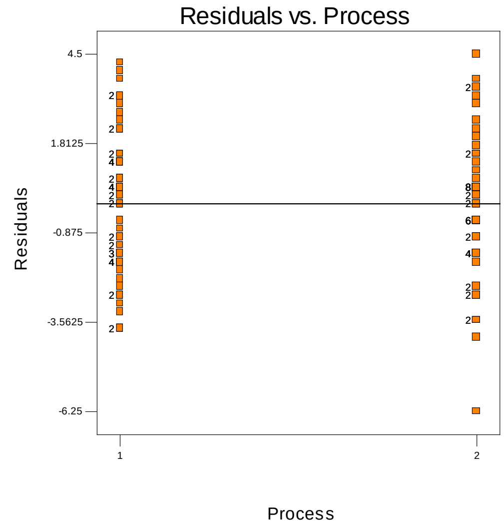
(b) Analyze the residuals from this experiment. Is there any indication that there is an outlier in one cell? If you find an outlier, remove it and repeat the analysis from part (a). What are your conclusions?

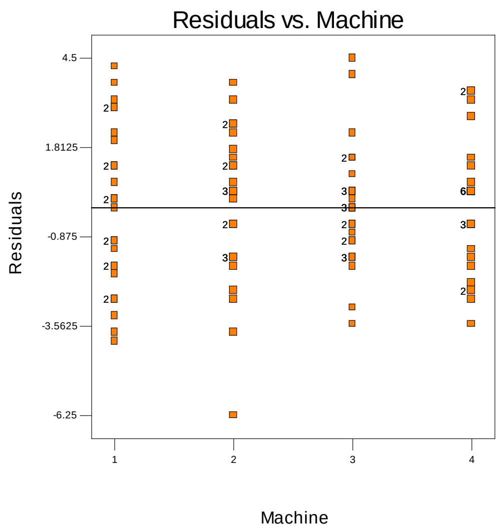
Standard Order 84, Time of Day at 3:00pm, Heat Treat #2, Machine #2, and length of 0, appears to be an outlier.











The following analysis was performed with the outlier described above removed. As with the original analysis, Heat Treat Process (*B*) and Machine (*C*) are significant, but now the *AC* interaction appears to be significant.

Design Expert Output

Response: Length

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	626.58	23	27.24	4.89	< 0.0001	significant
A	*19.17*	*2*	*9.59*	*1.72*	*0.1862*	
B	*85.39*	*1*	*85.39*	*15.33*	*0.0002*	
C	*411.89*	*3*	*137.30*	*24.65*	*< 0.0001*	
AB	*0.77*	*2*	*0.39*	*0.069*	*0.9331*	
AC	*81.55*	*6*	*13.59*	*2.44*	*0.0334*	
BC	*1.61*	*3*	*0.54*	*0.097*	*0.9617*	
ABC	*21.31*	*6*	*3.55*	*0.64*	*0.6996*	
Pure Error	395.42	71	5.57			
Cor Total	1022.00	94				

The Model F-value of 4.89 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

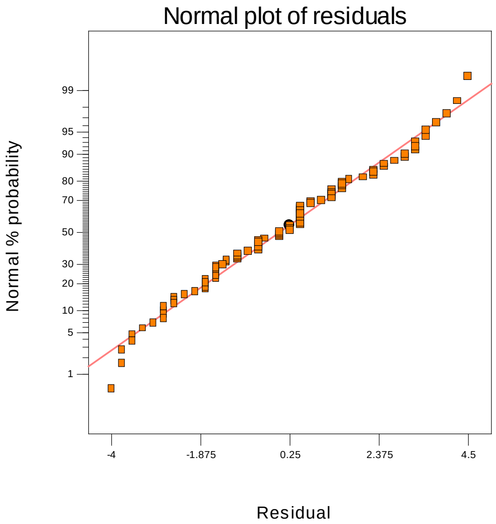
Std. Dev.	2.36	R-Squared	0.6131
Mean	4.00	Adj R-Squared	0.4878
C.V.	59.00	Pred R-Squared	0.3100
PRESS	705.17	Adeq Precision	7.447

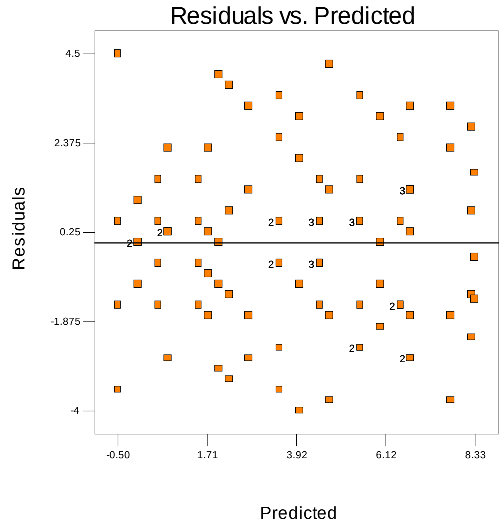
Term	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	4.05	1	0.24	3.56	4.53	
A[1]	-0.26	1	0.34	-0.95	0.42	
A[2]	-0.42	1	0.34	-1.10	0.26	
B-Process	-0.93	1	0.24	-1.42	-0.45	1.00
C[1]	-0.63	1	0.42	-1.46	0.21	
C[2]	2.18	1	0.43	1.33	3.03	
C[3]	-3.17	1	0.42	-4.00	-2.34	
A[1]B	0.090	1	0.34	-0.59	0.77	
A[2]B	-0.13	1	0.34	-0.81	0.55	
A[1]C[1]	0.60	1	0.59	-0.58	1.77	
A[2]C[1]	-1.50	1	0.59	-2.67	-0.32	
A[1]C[2]	-0.46	1	0.60	-1.65	0.73	
A[2]C[2]	-0.68	1	0.60	-1.87	0.51	
A[1]C[3]	0.26	1	0.59	-0.91	1.44	
A[2]C[3]	0.55	1	0.59	-0.63	1.72	
BC[1]	0.017	1	0.42	-0.82	0.85	
BC[2]	0.16	1	0.43	-0.69	1.01	
BC[3]	0.059	1	0.42	-0.77	0.89	
A[1]BC[1]	-0.17	1	0.59	-1.35	1.00	
A[2]BC[1]	0.30	1	0.59	-0.88	1.47	
A[1]BC[2]	-0.31	1	0.60	-1.50	0.88	
A[2]BC[2]	-0.72	1	0.60	-1.91	0.47	
A[1]BC[3]	-0.090	1	0.59	-1.27	1.09	
A[2]BC[3]	0.50	1	0.59	-0.67	1.68	

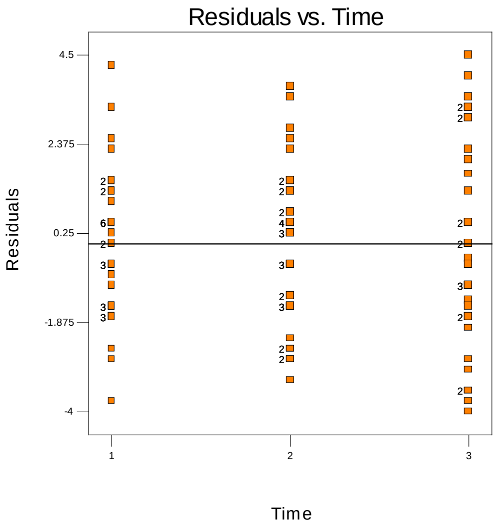
Final Equation in Terms of Coded Factors:

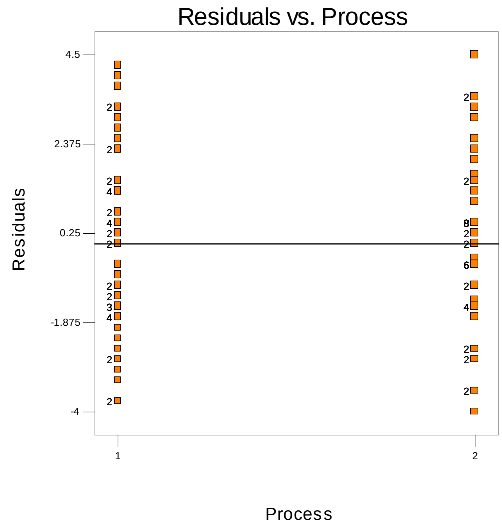
Length	=
+4.05	
-0.26	* A[1]
-0.42	* A[2]
-0.93	* B
-0.63	* C[1]
+2.18	* C[2]
-3.17	* C[3]
+0.090	* A[1]B
-0.13	* A[2]B
+0.60	* A[1]C[1]
-1.50	* A[2]C[1]
-0.46	* A[1]C[2]
-0.68	* A[2]C[2]
+0.26	* A[1]C[3]
+0.55	* A[2]C[3]
+0.017	* BC[1]
+0.16	* BC[2]
+0.059	* BC[3]
-0.17	* A[1]BC[1]
+0.30	* A[2]BC[1]
-0.31	* A[1]BC[2]
-0.72	* A[2]BC[2]
-0.090	* A[1]BC[3]
+0.50	* A[2]BC[3]

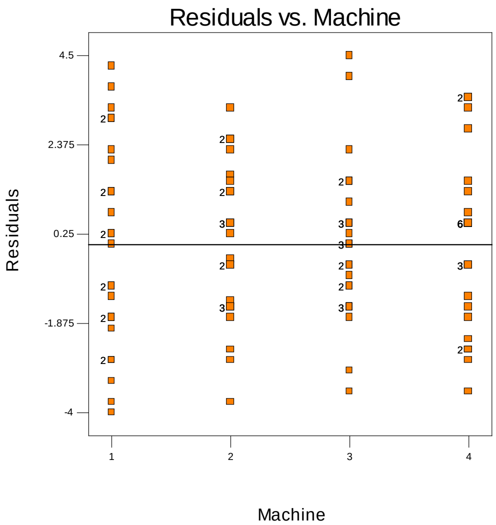
The following residual plots are acceptable. Both the normality and constant variance assumptions are satisfied


Normal plot of residuals
Normal % probability
99
95
90
80
70
50
30
20
10
5
1
-4
-1.875
0.25
2.375
4.5
Residual


Residuals vs. Predicted
Residuals
4.5
2.375
0.25
-1.875
-4
-0.50
1.71
3.92
6.12
8.33
Predicted


Residuals vs. Time
Residuals
4.5
2.375
0.25
-1.875
-4
1
2
3
Time


Residuals vs. Process
Residuals
4.5
2.375
0.25
-1.875
-4
1
2
Process


Residuals vs. Machine
Residuals
4.5
2.375
0.25
-1.875
-4
1
2
3
4
Machine

(c) Suppose that the observations in the cells are the lengths (coded) of bars processed together in heat treating and then cut sequentially (that is, in order) on the four machines. Analyze the data to determine the effects of the three factors on mean length.

The analysis with all effects and interactions included is shown below.

Design Expert Output

Response: Length
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	147.58	23	6.42		
A	*3.22*	*2*	*1.61*		
B	*25.01*	*1*	*25.01*		
C	*98.35*	*3*	*32.78*		
AB	*0.41*	*2*	*0.21*		
AC	*17.76*	*6*	*2.96*		
BC	*0.39*	*3*	*0.13*		
ABC	*2.44*	*6*	*0.41*		
Pure Error	0.000	0			
Cor Total	147.58	23			

By removing the three factor interaction from the model and applying it to the error, the analysis identifies factors *B* and *C* as being significant as well as the *AC* interaction.

Design Expert Output

Response: Length
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	145.14	17	8.54	20.97	0.0006	significant
A	*3.22*	*2*	*1.61*	*3.96*	*0.0801*	
B	*25.01*	*1*	*25.01*	*61.43*	*0.0002*	
C	*98.35*	*3*	*32.78*	*80.53*	*< 0.0001*	
AB	*0.41*	*2*	*0.21*	*0.51*	*0.6269*	
AC	*17.76*	*6*	*2.96*	*7.27*	*0.0146*	
BC	*0.39*	*3*	*0.13*	*0.32*	*0.8141*	
Residual	2.44	6	0.41			
Cor Total	147.58	23				

When removing the remaining insignificant effects from the model, factors *A*, *B*, *C*, and the *AC* interaction are significant.

Design Expert Output

Response: Avg
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	144.34	12	12.03	40.84	< 0.0001	significant
A	*3.22*	*2*	*1.61*	*5.47*	*0.0224*	
B	*25.01*	*1*	*25.01*	*84.92*	*< 0.0001*	
C	*98.35*	*3*	*32.78*	*111.32*	*< 0.0001*	
AC	*17.76*	*6*	*2.96*	*10.05*	*0.0006*	
Residual	3.24	11	0.29			
Cor Total	147.58	23				

The Model F-value of 40.84 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.54	R-Squared	0.9780
Mean	3.96	Adj R-Squared	0.9541

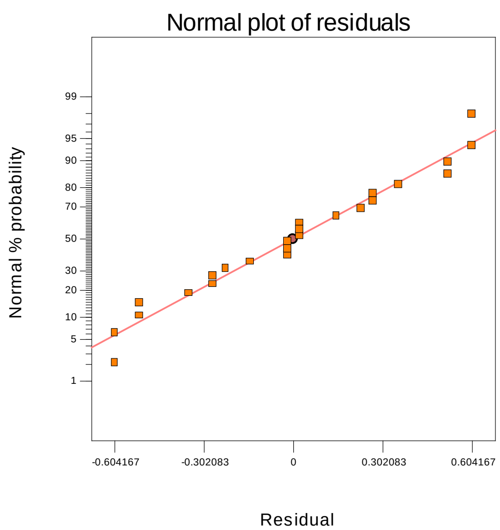
C.V.	13.71		Pred R-Squared	0.8955		
PRESS	15.42		Adeq Precision	20.760		

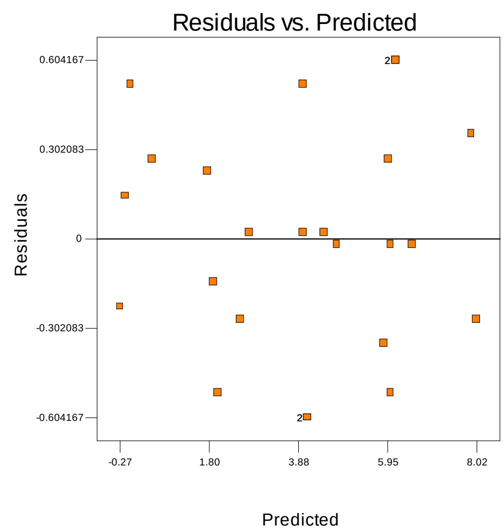
Term	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	3.96	1	0.11	3.71	4.20	
A[1]	-0.18	1	0.16	-0.52	0.17	
A[2]	-0.33	1	0.16	-0.68	0.011	
B-Process	-1.02	1	0.11	-1.26	-0.78	1.00
C[1]	-0.54	1	0.19	-0.96	-0.12	
C[2]	1.92	1	0.19	1.49	2.34	
C[3]	-3.08	1	0.19	-3.51	-2.66	
A[1]C[1]	0.51	1	0.27	-0.087	1.11	
A[2]C[1]	-1.58	1	0.27	-2.18	-0.99	
A[1]C[2]	-0.20	1	0.27	-0.80	0.40	
A[2]C[2]	-0.42	1	0.27	-1.01	0.18	
A[1]C[3]	0.18	1	0.27	-0.42	0.77	
A[2]C[3]	0.46	1	0.27	-0.14	1.06	

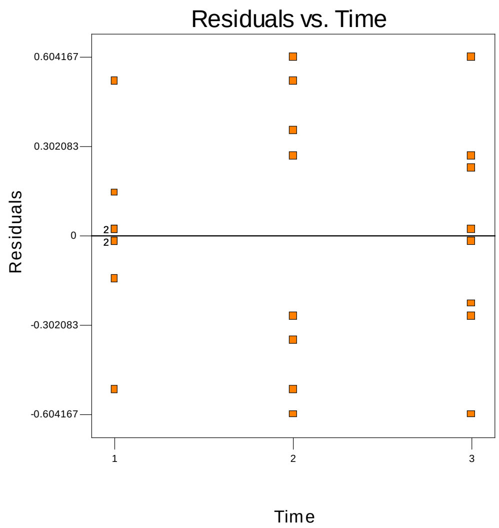
Final Equation in Terms of Coded Factors:

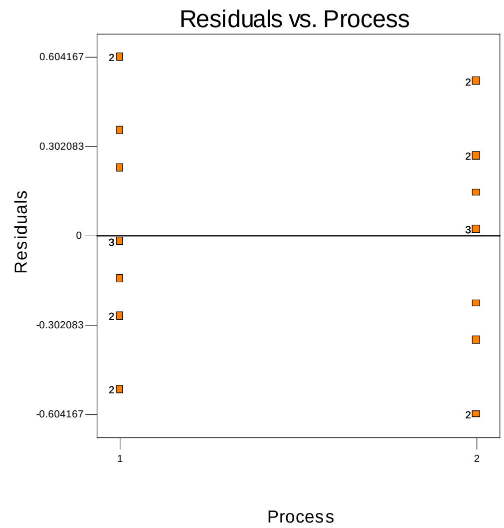
Avg	=
+3.96	
-0.18	* A[1]
-0.33	* A[2]
-1.02	* B
-0.54	* C[1]
+1.92	* C[2]
-3.08	* C[3]
+0.51	* A[1]C[1]
-1.58	* A[2]C[1]
-0.20	* A[1]C[2]
-0.42	* A[2]C[2]
+0.18	* A[1]C[3]
+0.46	* A[2]C[3]

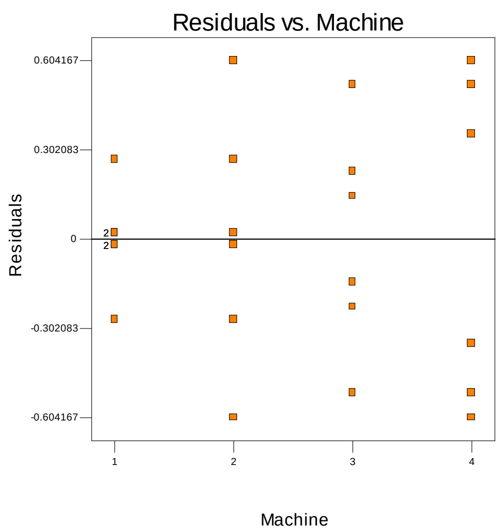
The following residual plots are acceptable. Both the normality and uniformity of variance assumptions are verified.











(d) Calculate the log variance of the observations in each cell. Analyze the average length and the log variance of the length for each of the 12 bars cut at each machine/heat treatment process combination. What conclusions can you draw?

Factors *A* and *C*, are significant as well as the *AB* and *BC* interactions. Factor *B* remains in the model for hierarchal purposes.

Design Expert Output

Response: Var Transform: Base 10 log Constant: 0

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	3.75	14	0.27	9.44	0.0010	significant
A	*1.45*	*2*	*0.72*	*25.45*	*0.0002*	
B	*0.093*	*1*	*0.093*	*3.27*	*0.1041*	
C	*0.59*	*3*	*0.20*	*6.95*	*0.0102*	
AB	*0.55*	*2*	*0.28*	*9.73*	*0.0056*	
AC	*1.07*	*6*	*0.18*	*6.28*	*0.0077*	
Residual	0.26	9	0.028			
Cor Total	4.01	23				

The Model F-value of 9.44 implies the model is significant. There is only a 0.10% chance that a "Model F-Value" this large could occur due to noise.

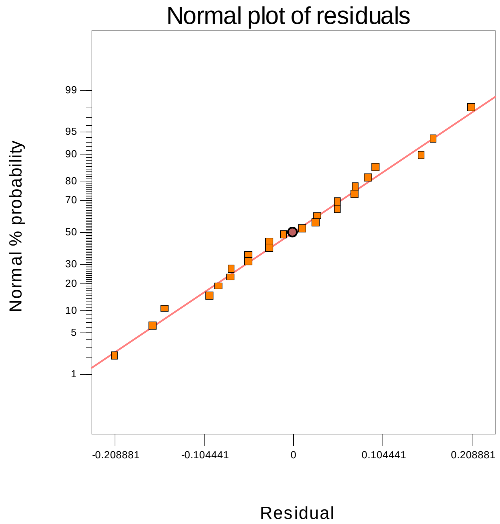
Std. Dev.	0.17	R-Squared	0.9363
Mean	0.65	Adj R-Squared	0.8371
C.V.	25.95	Pred R-Squared	0.5467
PRESS	1.82	Adeq Precision	11.415

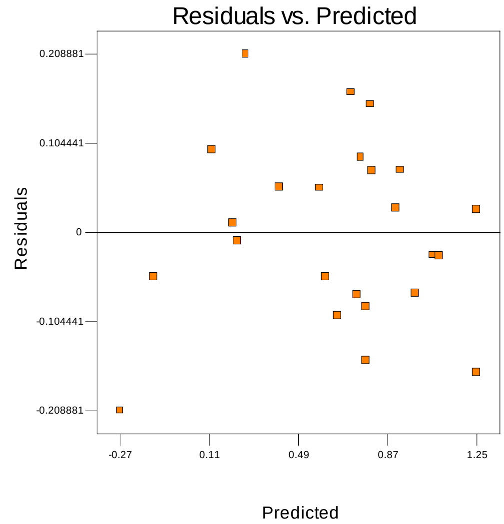
Term	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.65	1	0.034	0.57	0.73	
A[1]	-0.27	1	0.049	-0.38	-0.16	
A[2]	-0.055	1	0.049	-0.17	0.055	
B-Process	-0.062	1	0.034	-0.14	0.016	1.00
C[1]	0.22	1	0.060	0.085	0.35	
C[2]	0.066	1	0.060	-0.069	0.20	
C[3]	-0.19	1	0.060	-0.33	-0.058	
A[1]B	-0.21	1	0.049	-0.32	-0.096	
A[2]B	0.051	1	0.049	-0.059	0.16	
A[1]C[1]	0.38	1	0.084	0.19	0.57	
A[2]C[1]	-0.024	1	0.084	-0.22	0.17	
A[1]C[2]	-0.053	1	0.084	-0.24	0.14	
A[2]C[2]	-0.064	1	0.084	-0.25	0.13	
A[1]C[3]	-0.043	1	0.084	-0.23	0.15	
A[2]C[3]	-0.18	1	0.084	-0.37	0.012	

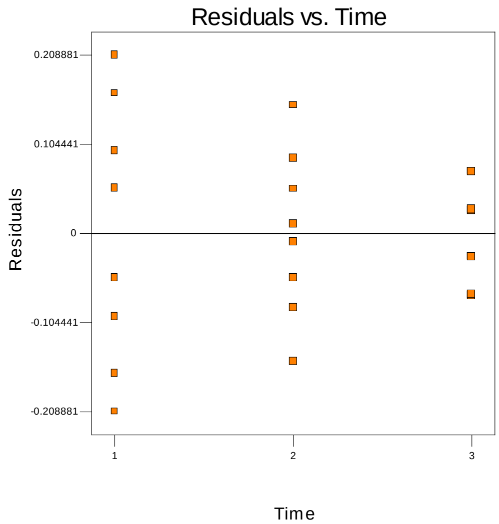
Final Equation in Terms of Coded Factors:

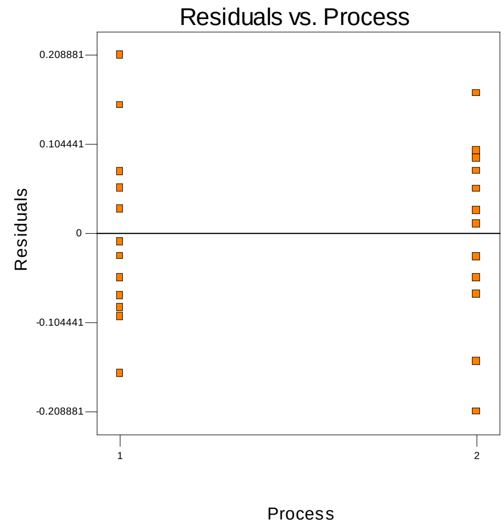
Log_{10}(Var)	=
+0.65	
-0.27	* A[1]
-0.055	* A[2]
-0.062	* B
+0.22	* C[1]
+0.066	* C[2]
-0.19	* C[3]
-0.21	* A[1]B
+0.051	* A[2]B
+0.38	* A[1]C[1]
-0.024	* A[2]C[1]
-0.053	* A[1]C[2]
-0.064	* A[2]C[2]
-0.043	* A[1]C[3]
-0.18	* A[2]C[3]

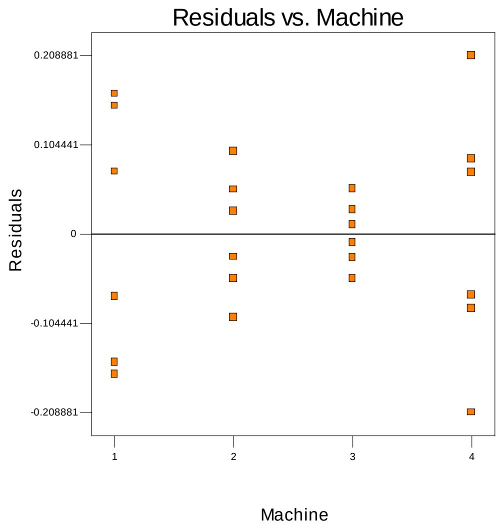
The following residual plots are acceptable, although the residuals versus time and residuals versus machine plots identify moderate inconsistency with variance.


Normal plot of residuals
Normal % probability
99
95
90
80
70
50
30
20
10
5
1
-0.208881
-0.104441
0
0.104441
0.208881
Residual


Residuals vs. Predicted
Residuals
0.208881
0.104441
0
-0.104441
-0.208881
-0.27
0.11
0.49
0.87
1.25
Predicted


Residuals vs. Time
Residuals
0.208881
0.104441
0
-0.104441
-0.208881
1
2
3
Time


Residuals vs. Process
Residuals
0.208881
0.104441
0
-0.104441
-0.208881
1
2
Process


Residuals vs. Machine
Residuals
0.208881
0.104441
0
-0.104441
-0.208881
1
2
3
4
Machine

(e) Suppose the time at which a bar is cut really cannot be controlled during routine production. Analyze the average length and the log variance of the length for each of the 12 bars cut at each machine/heat treatment process combination. What conclusions can you draw?

The analysis of the average length is as follows:

Design Expert Output

Response: Avg
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	41.25	7	5.89		
A	8.34	1	8.34		
B	32.78	3	10.93		
AB	0.13	3	0.043		
Pure Error	0.000	0			
Cor Total	41.25	7			

Because the Means Square of the AB interaction is much less than the main effects, it is removed from the model and placed in the error. The average length is significantly affected by factor *A*, Heat Treat Process, and factor *B*, Machine.

Design Expert Output

Response: Avg
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	41.12	4	10.28	240.06	0.0004	significant
A	*8.34*	*1*	*8.34*	*194.68*	*0.0008*	
B	*32.78*	*3*	*10.93*	*255.19*	*0.0004*	
Residual	0.13	3	0.043			
Cor Total	41.25	7				

The Model F-value of 240.06 implies the model is significant. There is only a 0.04% chance that a "Model F-Value" this large could occur due to noise.

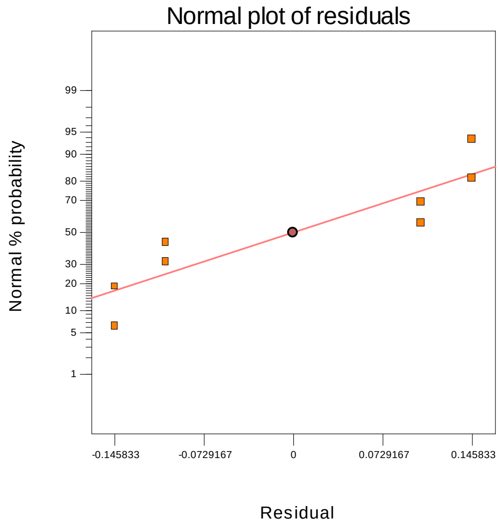
Std. Dev.	0.21	R-Squared	0.9969
Mean	3.96	Adj R-Squared	0.9927
C.V.	5.23	Pred R-Squared	0.9779
PRESS	0.91	Adeq Precision	43.042

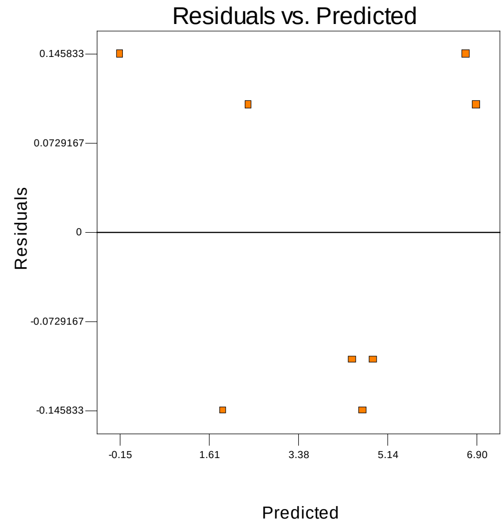
Term	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	3.96	1	0.073	3.73	4.19	
A-Process	-1.02	1	0.073	-1.25	-0.79	1.00
B[1]	-0.54	1	0.13	-0.94	-0.14	
B[2]	1.92	1	0.13	1.51	2.32	
B[3]	-3.08	1	0.13	-3.49	-2.68	

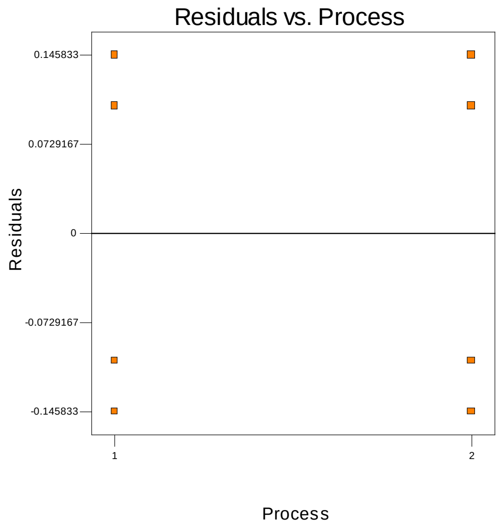
Final Equation in Terms of Coded Factors:

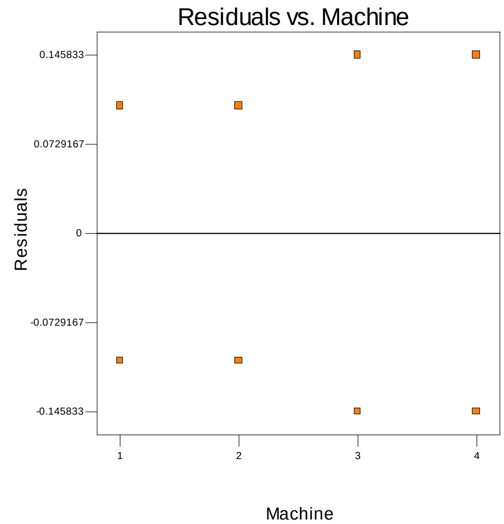
Avg =
+3.96
-1.02 * A
-0.54 * B[1]
+1.92 * B[2]
-3.08 * B[3]

The following residual plots are acceptable. Both the normality and uniformity of variance assumptions are verified.









The Log(Var) analysis is shown below.

Design Expert Output

Response: Log(Var)
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	0.16	7	0.023		
A	*1.070E-003*	*1*	*1.070E-003*		
B	*0.14*	*3*	*0.046*		
AB	*0.023*	*3*	*7.755E-003*		
Pure Error	0.000	0			
Cor Total	0.16	7			

Because the factor *A* and the *AB* interaction small Mean Squares, they were removed from the model and placed in the error. From the following analysis of variance, Machine (*B*) is significant.

Design Expert Output

Response: Var Transform: Base 10 log Constant: 0
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	0.14	3	0.046	7.58	0.0398	significant
B	*0.14*	*3*	*0.046*	*7.58*	*0.0398*	
Residual	0.024	4	6.084E-003			
Cor Total	0.16	7				

The Model F-value of 7.58 implies the model is significant. There is only a 3.98% chance that a "Model F-Value" this large could occur due to noise.

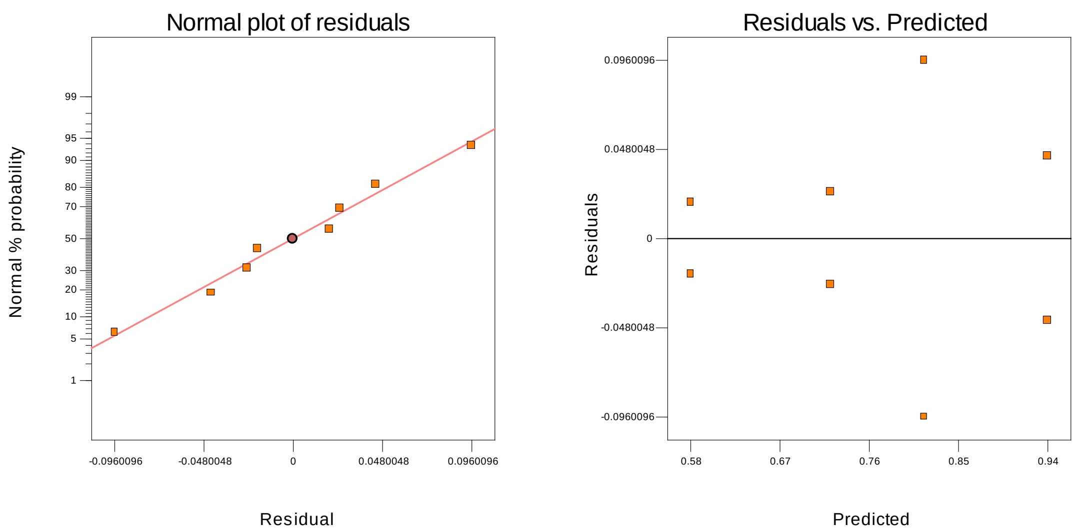
Std. Dev.	0.078	R-Squared	0.8504
Mean	0.77	Adj R-Squared	0.7383
C.V.	10.17	Pred R-Squared	0.4018
PRESS	0.097	Adeq Precision	6.521

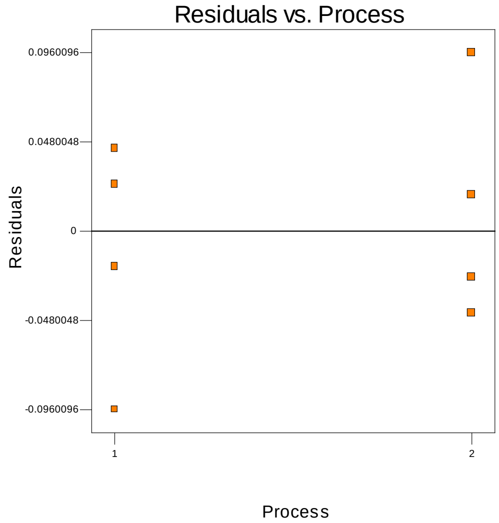
Term	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.77	1	0.028	0.69	0.84	
B[1]	0.18	1	0.048	0.043	0.31	
B[2]	0.051	1	0.048	-0.081	0.18	
B[3]	-0.18	1	0.048	-0.32	-0.051	

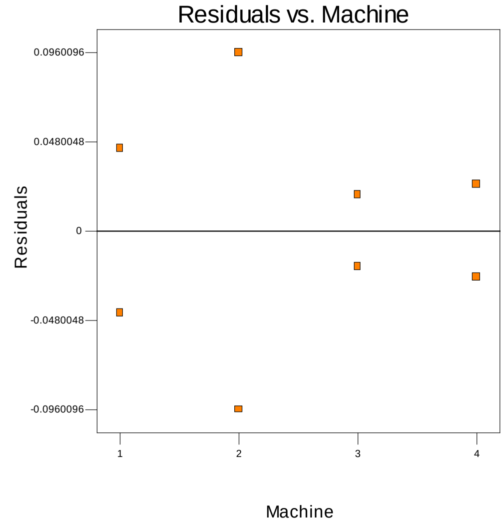
Final Equation in Terms of Coded Factors:

$\text{Log}_{10}(\text{Var})$ =
+0.77
+0.18 * B[1]
+0.051 * B[2]
-0.18 * B[3]

The following residual plots are acceptable.







9.24. Reconsider the experiment in Problem 9.23. Suppose that it was necessary to estimate all main effects and two-factor interactions, but the full factorial with 24 runs (not counting replication) was too expensive. Recommend an alternative design.

The design below is a D-optimal design generated by Design Expert software. For this design, zero additional runs were selected for lack of fit and replication. This is the minimum number of runs required to estimate all main effects and two-factor interactions. Additional runs could have been selected for lack of fit or replication.

Run	Heat Treatment	Machine	Time of Day
1	2	4	3
2	1	2	2
3	1	3	3
4	2	2	2
5	1	1	1
6	2	3	3
7	2	2	1
8	2	4	2
9	1	2	3
10	1	2	1
11	2	1	2
12	2	1	3
13	1	4	1
14	2	3	1
15	1	1	2
16	1	1	3
17	2	4	1
18	1	3	2

9.25. Suppose there are four three-level categorical factors and a style two-level continuous factor. What is the minimum number of runs required to estimate all main effects and two-factor interactions? Construct this design.

A minimum of 42 runs are required to estimate all main effects and two factor interactions. Below is a D-optimal design with 42 runs.

Run	Factor *A*	Factor *B*	Factor *C*	Factor *D*	Factor *E*
1	2	1	1	1	-1
2	1	1	2	1	+1
3	3	3	2	3	-1
4	2	2	1	3	-1
5	1	1	1	2	+1
6	3	2	3	3	-1
7	1	2	1	2	-1
8	1	1	3	2	+1
9	2	1	3	2	+1
10	1	2	2	2	+1
11	2	3	3	1	-1
12	3	2	2	3	+1
13	2	1	3	1	-1
14	3	1	1	2	+1
15	2	3	1	2	-1
16	3	1	3	3	+1
17	2	2	2	1	-1
18	3	3	1	2	+1
19	1	1	3	3	-1
20	3	3	3	2	-1
21	3	2	1	1	-1
22	1	3	3	3	-1
23	1	3	2	3	+1
24	2	1	2	2	-1
25	1	3	2	2	-1
26	3	2	3	1	+1
27	1	3	1	1	-1
28	2	2	3	3	+1
29	1	3	3	1	+1
30	3	1	1	3	-1
31	2	3	2	2	+1
32	1	2	3	1	-1
33	3	2	2	2	-1
34	2	3	1	1	+1
35	3	1	2	1	-1
36	3	3	2	1	+1
37	1	2	1	3	+1
38	2	1	2	3	-1
39	3	1	1	1	+1
40	2	3	3	3	+1
41	1	2	2	3	-1
42	2	2	1	2	+1

9.26. Reconsider the experiment in Problem 9.25. Construct a design with N = 48 runs and compare it to the design you constructed in Problem 9.25.

When generating the D-optimal design in Design Expert software, an additional six runs were included for lack of fit. The resulting design is shown below. As with the design created in Problem 9.25, the three level factors are balanced with equal runs at each individual factor level, and the two level factor is very slightly unbalanced with 25 low level runs and 23 high level runs.

Run	Factor *A*	Factor *B*	Factor *C*	Factor *D*	Factor *E*
1	1	1	3	3	-1
2	3	3	1	1	-1
3	2	3	3	3	+1
4	3	1	2	1	-1
5	1	2	3	2	+1
6	2	1	2	3	-1
7	2	2	3	3	-1
8	3	1	2	2	+1
9	2	2	1	1	+1
10	1	3	3	3	+1
11	2	2	1	2	-1
12	3	2	2	3	-1
13	1	3	1	2	-1
14	2	3	1	2	+1
15	3	1	3	1	-1
16	1	2	2	3	+1
17	3	3	3	3	-1
18	1	2	2	1	+1
19	1	1	1	2	+1
20	3	3	2	3	+1
21	2	2	3	1	-1
22	3	3	3	2	+1
23	2	1	1	1	-1
24	1	2	1	1	-1
25	3	1	3	3	+1
26	1	3	2	3	-1
27	3	1	1	1	+1
28	1	3	3	1	-1
29	2	1	3	2	-1
30	1	3	2	2	+1
31	3	2	3	2	-1
32	2	3	2	1	+1
33	1	1	2	2	-1
34	2	1	1	3	+1
35	2	2	2	2	+1
36	2	3	2	2	-1
37	3	1	1	3	-1
38	1	1	2	1	+1
39	3	2	1	2	+1
40	1	3	1	1	+1
41	2	3	1	3	-1
42	3	2	3	1	+1
43	2	1	3	1	+1
44	1	2	1	2	-1
45	3	2	2	3	+1
46	3	1	1	1	-1
47	1	2	2	2	-1
48	2	3	3	3	-1

9.27. Reconsider the experiment in Problem 9.25. Suppose that you are only interested in main effects. Construct a design with $N = 12$ runs for this experiment.

Run	Factor *A*	Factor *B*	Factor *C*	Factor *D*	Factor *E*
1	1	3	2	3	+1
2	1	2	3	2	-1
3	2	2	1	3	+1
4	1	1	1	1	-1
5	3	2	2	1	-1
6	3	1	3	3	-1
7	3	1	1	2	+1
8	3	3	3	1	+1
9	2	1	3	1	+1
10	2	3	2	2	-1
11	2	2	1	2	-1
12	1	3	2	1	-1

9.28. An article in the *Journal of Chemical Technology and Biotechnology* ("A Study of Antifungal Antibiotic Production by *Thermomomospora* sp MTCC 3340 Using Full Factorial Design," 2003, Vol. 78, pp. 605-610) investigated three independent variables – concentration of carbon source (glucose), concentration of nitrogen source (soybean meal), and temperature of incubation for their effects on the production of antifungal antibiotic by the isolate *Thermomonospora* sp MTCC 3340. A 3^3 factorial design was conducted and the results are shown in the table on the previous page.

Run	% Carbon	% Nitrogen	Temp	Activity
1	2	0.5	25	25.84
2	2	1	25	51.86
3	2	3	25	32.59
4	5	0.5	25	20.48
5	5	1	25	25.84
6	5	3	25	12.87
7	7.5	0.5	25	20.48
8	7.5	1	25	25.84
9	7.5	3	25	10.2
10	2	0.5	30	51.86
11	2	1	30	131.33
12	2	3	30	41.11
13	5	0.5	30	41.11
14	5	1	30	104.11
15	5	3	30	32.59
16	7.5	0.5	30	65.42
17	7.5	1	30	82.53
18	7.5	3	30	51.86
19	2	0.5	35	41.11
20	2	1	35	104.11
21	2	3	35	32.59
22	5	0.5	35	32.59

23	5	1	35	82.53
24	5	3	35	25.84
25	7.5	0.5	35	51.86
26	7.5	1	35	65.42
27	7.5	3	35	41.11

(a) Analyze the data from this experiment.

The ANOVA is shown below. Factor *B*, and the B^2 and C^2 quadratic terms appear to be significant.

Design Expert Output

Response 1 Activity
ANOVA for Response Surface Quadratic 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	19338.43	9	2148.71	8.63	< 0.0001	significant
A-Carbon	*433.77*	*1*	*433.77*	*1.74*	*0.2043*	
B-Nitrogen	*280.10*	*1*	*280.10*	*1.13*	*0.3036*	
C-Temperature	*2913.48*	*1*	*2913.48*	*11.70*	*0.0033*	
AB	*50.78*	*1*	*50.78*	*0.20*	*0.6572*	
AC	*97.05*	*1*	*97.05*	*0.39*	*0.5407*	
BC	*215.04*	*1*	*215.04*	*0.86*	*0.3657*	
A2	*486.58*	*1*	*486.58*	*1.95*	*0.1801*	
B2	*7503.07*	*1*	*7503.07*	*30.14*	*< 0.0001*	
C2	*4642.23*	*1*	*4642.23*	*18.65*	*0.0005*	
Residual	4231.91	17	248.94			
Cor Total	23570.33	26				

Std. Dev.	15.78	R-Squared	0.8205
Mean	48.34	Adj R-Squared	0.7254
C.V. %	32.64	Pred R-Squared	0.5795
PRESS	9911.08	Adeq Precision	10.086

(b) Fit a second-order model to the activity response. Construct contour plots and response surface plots to assist in interpreting the results of this experiment.

The model, contour plot, and response surface plot are shown below. Because factor *A* was not significant, the plots are shown with factors *B* and *C* only.

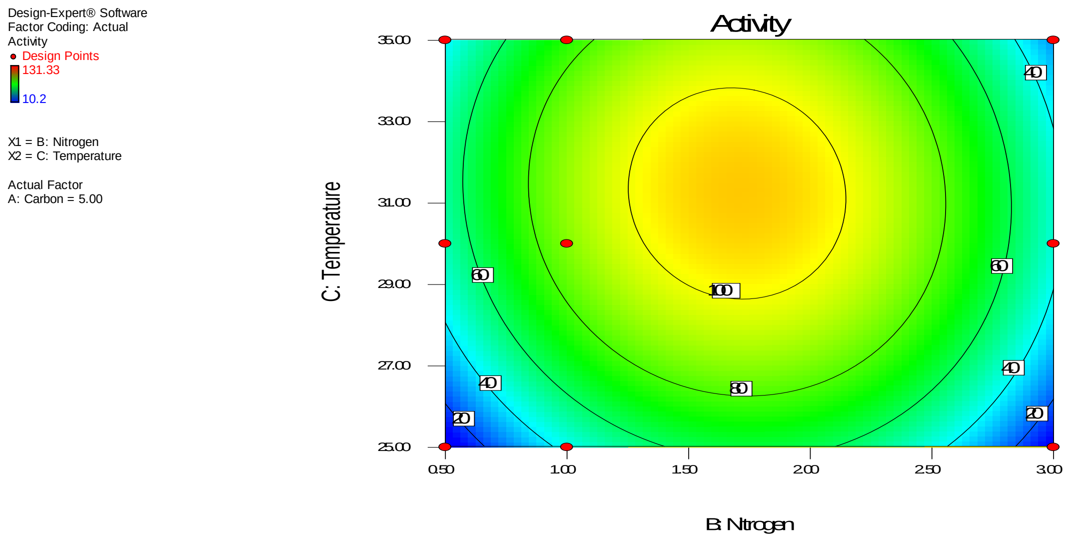
Design Expert Output

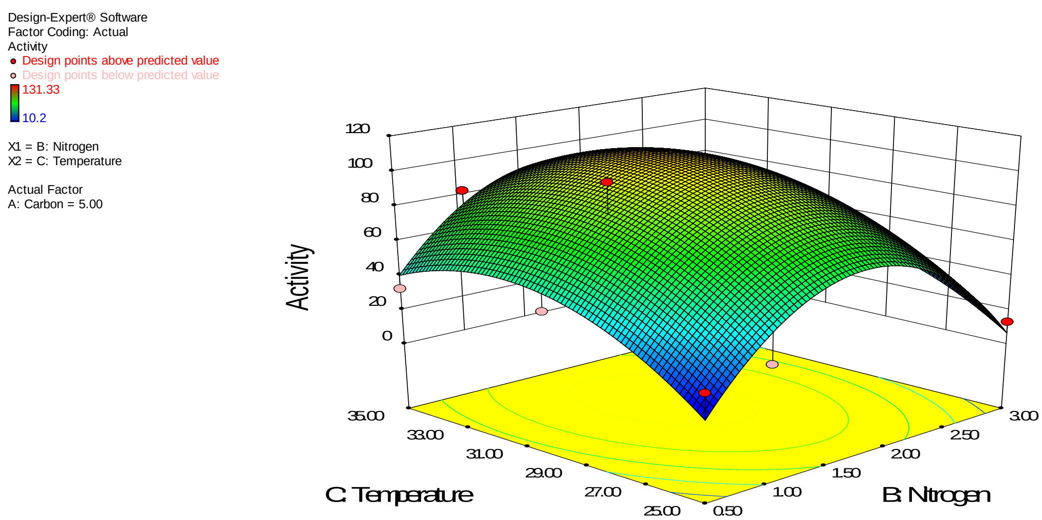
Term	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	106.17	1	11.02	82.91	129.43	
A-Carbon	-5.04	1	3.82	-13.09	3.01	1.06
B-Nitrogen	-3.95	1	3.72	-11.80	3.90	1.12
C-Temperature	13.07	1	3.82	5.01	21.13	1.05
AB	1.94	1	4.30	-7.13	11.01	1.05
AC	2.84	1	4.55	-6.76	12.44	1.00
BC	-4.00	1	4.30	-13.08	5.08	1.05
A2	9.09	1	6.50	-4.63	22.81	1.00
B2	-58.48	1	10.65	-80.95	-36.00	1.12
C2	-27.82	1	6.44	-41.41	-14.23	1.00

Final Equation in Terms of Actual Factors:

Activity =
-1046.32378

-20.43903	* Carbon
+144.34481	* Nitrogen
+69.50973	* Temperature
+0.56468	* Carbon * Nitrogen
+0.20654	* Carbon * Temperature
-0.64000	* Nitrogen * Temperature
+1.20237	* Carbon2
-37.42422	* Nitrogen2
-1.11262	* Temperature2





(c) What operating conditions would you recommend to optimize this process?

Set factor *B*, Nitrogen, at approximately 1.7, and factor *C*, Temperature, at 31, maximizes the Activity. Factor *A* is not significant, so any level can be chosen depending on the manufacturing preferences.

9.29. Construct a minimum run D-optimal resolution IV design for 10 factors. Find the alias relationships. What approach would you recommend for analyzing the data from this experiment?

A	B	C	D	E	F	G	H	J	K
-	+	+	-	-	-	+	+	-	-
+	+	+	+	+	-	+	+	+	+
-	+	-	-	+	+	-	+	+	-
-	+	-	+	-	-	-	+	-	+
+	+	-	-	-	-	-	-	+	-
+	+	-	-	+	+	+	-	-	+
-	-	-	-	+	-	+	-	+	+
+	-	-	+	+	-	+	+	-	-
-	+	+	+	+	-	-	-	-	-
-	+	-	+	-	+	+	-	+	-
+	-	-	+	+	+	-	-	+	+
-	-	-	-	-	+	-	-	-	-
+	-	+	+	-	-	+	-	-	+
+	-	+	-	+	+	+	-	+	-
-	-	+	+	+	+	+	+	-	+
-	-	+	+	-	-	-	+	+	-
+	+	+	+	-	+	-	+	-	-
-	+	+	-	-	+	-	-	+	+
+	-	-	-	-	+	+	+	+	+
+	-	+	-	+	-	-	+	-	+

The alias relationships for the main effects and two factor interactions are as follows:

[Intercept] = Intercept - 0.333 * BE - 0.333 * BG - 0.333 * BK + 0.333 * CD
+ 0.333 * CH - 0.333 * CJ + 0.333 * DH - 0.333 * DJ + 0.333 * EG + 0.333 * EK
+ 0.333 * GK - 0.333 * HJ

[A] = A	[B] = B	**[C] = C**	[D] = D	[E] = E
[F] = F	[G] = G	[H] = H	[J] = J	[K] = K

[AB] = AB + 0.333 * BE + 0.333 * BG + 0.333 * BK + 0.667 * CD + 0.667 * CH
+ 0.333 * CJ + 0.667 * DH + 0.333 * DJ + 0.667 * EG + 0.667 * EK - FJ
+ 0.667 * GK + 0.333 * HJ

[AC] = AC - 0.333 * BC + 0.667 * BD + 0.667 * BH + 0.333 * CE + 0.333 * CG
- 0.667 * DE + 0.333 * DK + 0.667 * EJ - FK - 0.667 * GH + 0.667 * GJ
+ 0.333 * HK - 0.333 * JK

[AD] = AD + 0.667 * BC - 0.333 * BD + 0.667 * BH - 0.667 * CE + 0.333 * CG
+ 0.333 * DE + 0.333 * DK + 0.667 * EJ - FG + 0.333 * GH - 0.333 * GJ
- 0.667 * HK + 0.667 * JK

[AE] = AE - 0.333 * BE + 0.667 * BG + 0.667 * BK - 0.667 * CD + 0.333 * CH
+ 0.667 * CJ + 0.333 * DH + 0.667 * DJ + 0.333 * EG + 0.333 * EK - FH
- 0.667 * GK - 0.333 * HJ

[AF] = AF + 0.333 * BC + 0.333 * BD + 0.333 * BH - BJ - 0.333 * CE - 0.333 * CG
- CK - 0.333 * DE - DG - 0.333 * DK - EH + 0.333 * EJ - 0.333 * GH + 0.333 * GJ
- 0.333 * HK + 0.333 * JK

[AG] = AG + 0.667 * BE - 0.333 * BG + 0.667 * BK + 0.333 * CD - 0.667 * CH
+ 0.667 * CJ - DF + 0.333 * DH - 0.333 * DJ + 0.333 * EG - 0.667 * EK
+ 0.333 * GK + 0.667 * HJ

[AH] = AH + 0.667 * BC + 0.667 * BD - 0.333 * BH + 0.333 * CE - 0.667 * CG
+ 0.333 * DE - 0.667 * DK - EF - 0.333 * EJ + 0.333 * GH + 0.667 * GJ
+ 0.333 * HK + 0.667 * JK

[AJ] = AJ + 0.333 * BC + 0.333 * BD - BF + 0.333 * BH + 0.667 * CE + 0.667 * CG
+ 0.667 * DE + 0.667 * DK + 0.333 * EJ + 0.667 * GH + 0.333 * GJ + 0.667 * HK
+ 0.333 * JK

[AK] = AK + 0.667 * BE + 0.667 * BG - 0.333 * BK + 0.333 * CD - CF + 0.333 * CH
- 0.333 * CJ - 0.667 * DH + 0.667 * DJ - 0.667 * EG + 0.333 * EK + 0.333 * GK
+ 0.667 * HJ

Forward selection regression should be used for analyzing the experiment. Hierarchy should be considered.

9.30. Construct a minimum run D-optimal resolution IV design for 12 factors. Find the alias relationships. What approach would you recommend for analyzing the data from this experiment?

A	B	C	D	E	F	G	H	J	K	L	M
-	-	+	-	-	+	+	-	-	-	-	+
-	-	-	-	+	+	-	-	+	+	+	+
-	+	-	+	+	+	+	+	-	+	-	+
-	+	-	-	-	-	-	+	+	-	-	+
+	+	-	+	-	+	+	-	+	-	+	+
-	+	+	-	-	+	+	+	+	+	+	-
-	-	-	+	-	+	-	+	-	-	+	-
-	-	+	+	+	-	+	+	+	-	+	+
+	-	-	-	+	+	+	+	+	-	-	-
-	-	-	+	-	-	+	-	+	+	-	-
+	-	-	-	-	-	+	+	-	+	+	+
-	+	-	-	+	-	+	-	-	-	+	-
+	+	-	+	+	-	-	+	+	+	+	-
+	+	+	+	-	-	+	+	-	-	-	-
+	-	+	-	-	-	-	-	+	-	+	-
+	-	+	+	+	+	+	-	-	+	+	-
-	-	+	-	+	-	-	+	-	+	-	-
+	-	-	+	+	-	-	-	-	-	-	+
+	+	+	-	+	-	+	-	+	+	-	+
+	+	-	-	-	+	-	-	-	+	-	-
-	+	+	+	-	-	-	-	-	+	+	+
+	+	+	-	+	+	-	+	-	-	+	+
-	+	+	+	+	+	-	-	+	-	-	-
+	-	+	+	-	+	-	+	+	+	-	+

The alias relationships for the main effects and two factor interactions are as follows:

[Intercept] = Intercept

[A] = A	[B] = B	[C] = C	[D] = D	[E] = E
[F] = F	[G] = G	[H] = H	[J] = J	[K] = K
[L] = L	[M] = M			

[AB] = AB - 0.333 * CD + 0.333 * CE - 0.333 * CF + 0.333 * CG + 0.333 * CH - 0.333 * CJ - 0.333 * CK - 0.333 * CL + 0.333 * CM - 0.333 * DE - 0.333 * DF + 0.333 * DG + 0.333 * DH + 0.333 * DJ - 0.333 * DK + 0.333 * DL - 0.333 * DM - 0.333 * EF - 0.333 * EG + 0.333 * EH + 0.333 * EJ + 0.333 * EK + 0.333 * EL + 0.333 * EM - 0.333 * FG - 0.333 * FH - 0.333 * FJ - 0.333 * FK + 0.333 * FL + 0.333 * FM - 0.333 * GH + 0.333 * GJ - 0.333 * GK - 0.333 * GL + 0.333 * GM - 0.333 * HJ - 0.333 * HK + 0.333 * HL - 0.333 * HM + 0.333 * JK + 0.333 * JL + 0.333 * JM - 0.333 * KL - 0.333 * KM + 0.333 * LM

[AC] = AC - 0.333 * BD + 0.333 * BE - 0.333 * BF + 0.333 * BG + 0.333 * BH - 0.333 * BJ - 0.333 * BK - 0.333 * BL + 0.333 * BM - 0.333 * DE + 0.333 * DF + 0.333 * DG + 0.333 * DH - 0.333 * DJ + 0.333 * DK - 0.333 * DL - 0.333 * DM + 0.333 * EF + 0.333 * EG - 0.333 * EH - 0.333 * EJ + 0.333 * EK + 0.333 * EL + 0.333 * EM - 0.333 * FG + 0.333 * FH - 0.333 * FJ + 0.333 * FK + 0.333 * FL + 0.333 * FM - 0.333 * GH - 0.333 * GJ + 0.333 * GK - 0.333 * GL - 0.333 * GM - 0.333 * HJ - 0.333 * HK - 0.333 * HL + 0.333 * HM + 0.333 * JK - 0.333 * JL + 0.333 * JM - 0.333 * KL + 0.333 * KM - 0.333 * LM

[AD] = AD - 0.333 * BC - 0.333 * BE - 0.333 * BF + 0.333 * BG + 0.333 * BH + 0.333 * BJ - 0.333 * BK + 0.333 * BL - 0.333 * BM - 0.333 * CE + 0.333 * CF + 0.333 * CG + 0.333 * CH - 0.333 * CJ + 0.333 * CK - 0.333 * CL - 0.333 * CM - 0.333 * EF - 0.333 * EG - 0.333 * EH - 0.333 * EJ + 0.333 * EK + 0.333 * EL - 0.333 * EM + 0.333 * FG - 0.333 * FH + 0.333 * FJ + 0.333 * FK + 0.333 * FL + 0.333 * FM - 0.333 * GH - 0.333 * GJ - 0.333 * GK + 0.333 * GL - 0.333 * GM + 0.333 * HJ + 0.333 * HK - 0.333 * HL - 0.333 * HM + 0.333 * JK + 0.333 * JL + 0.333 * JM + 0.333 * KL - 0.333 * KM - 0.333 * LM

[AE] = AE + 0.333 * BC - 0.333 * BD - 0.333 * BF - 0.333 * BG + 0.333 * BH + 0.333 * BJ + 0.333 * BK + 0.333 * BL + 0.333 * BM - 0.333 * CD + 0.333 * CF + 0.333 * CG - 0.333 * CH - 0.333 * CJ + 0.333 * CK + 0.333 * CL + 0.333 * CM - 0.333 * DF - 0.333 * DG - 0.333 * DH - 0.333 * DJ + 0.333 * DK + 0.333 * DL - 0.333 * DM + 0.333 * FG + 0.333 * FH - 0.333 * FJ - 0.333 * FK + 0.333 * FL - 0.333 * FM - 0.333 * GH + 0.333 * GJ + 0.333 * GK - 0.333 * GL - 0.333 * GM + 0.333 * HJ - 0.333 * HK + 0.333 * HL - 0.333 * HM + 0.333 * JK - 0.333 * JL - 0.333 * JM + 0.333 * KL - 0.333 * KM - 0.333 * LM

[AF] = AF - 0.333 * BC - 0.333 * BD - 0.333 * BE - 0.333 * BG - 0.333 * BH - 0.333 * BJ - 0.333 * BK + 0.333 * BL + 0.333 * BM + 0.333 * CD + 0.333 * CE - 0.333 * CG + 0.333 * CH - 0.333 * CJ + 0.333 * CK + 0.333 * CL + 0.333 * CM - 0.333 * DE + 0.333 * DG - 0.333 * DH + 0.333 * DJ + 0.333 * DK + 0.333 * DL + 0.333 * DM + 0.333 * EG + 0.333 * EH - 0.333 * EJ - 0.333 * EK + 0.333 * EL - 0.333 * EM - 0.333 * GH + 0.333 * GJ - 0.333 * GK + 0.333 * GL - 0.333 * GM + 0.333 * HJ - 0.333 * HK - 0.333 * HL + 0.333 * HM - 0.333 * JK - 0.333 * JL + 0.333 * JM - 0.333 * KL - 0.333 * KM + 0.333 * LM

[AG] = AG + 0.333 * BC + 0.333 * BD - 0.333 * BE - 0.333 * BF - 0.333 * BH + 0.333 * BJ - 0.333 * BK - 0.333 * BL + 0.333 * BM + 0.333 * CD + 0.333 * CE - 0.333 * CF - 0.333 * CH - 0.333 * CJ + 0.333 * CK - 0.333 * CL - 0.333 * CM

- 0.333 * DE + 0.333 * DF - 0.333 * DH - 0.333 * DJ - 0.333 * DK + 0.333 * DL
- 0.333 * DM + 0.333 * EF - 0.333 * EH + 0.333 * EJ + 0.333 * EK - 0.333 * EL
- 0.333 * EM - 0.333 * FH + 0.333 * FJ - 0.333 * FK + 0.333 * FL - 0.333 * FM
- 0.333 * HJ - 0.333 * HK - 0.333 * HL - 0.333 * HM - 0.333 * JK - 0.333 * JL
+ 0.333 * JM + 0.333 * KL + 0.333 * KM + 0.333 * LM

[AH] = AH + 0.333 * BC + 0.333 * BD + 0.333 * BE - 0.333 * BF - 0.333 * BG
- 0.333 * BJ - 0.333 * BK + 0.333 * BL - 0.333 * BM + 0.333 * CD - 0.333 * CE
+ 0.333 * CF - 0.333 * CG - 0.333 * CJ - 0.333 * CK - 0.333 * CL + 0.333 * CM
- 0.333 * DE - 0.333 * DF - 0.333 * DG + 0.333 * DJ + 0.333 * DK - 0.333 * DL
- 0.333 * DM + 0.333 * EF - 0.333 * EG + 0.333 * EJ - 0.333 * EK + 0.333 * EL
- 0.333 * EM - 0.333 * FG + 0.333 * FJ - 0.333 * FK - 0.333 * FL + 0.333 * FM
- 0.333 * GJ - 0.333 * GK - 0.333 * GL - 0.333 * GM + 0.333 * JK - 0.333 * JL
- 0.333 * JM + 0.333 * KL + 0.333 * KM + 0.333 * LM

[AJ] = AJ - 0.333 * BC + 0.333 * BD + 0.333 * BE - 0.333 * BF + 0.333 * BG
- 0.333 * BH + 0.333 * BK + 0.333 * BL + 0.333 * BM - 0.333 * CD - 0.333 * CE
- 0.333 * CF - 0.333 * CG - 0.333 * CH + 0.333 * CK - 0.333 * CL + 0.333 * CM
- 0.333 * DE + 0.333 * DF - 0.333 * DG + 0.333 * DH + 0.333 * DK + 0.333 * DL
+ 0.333 * DM - 0.333 * EF + 0.333 * EG + 0.333 * EH + 0.333 * EK - 0.333 * EL
- 0.333 * EM + 0.333 * FG + 0.333 * FH - 0.333 * FK - 0.333 * FL + 0.333 * FM
- 0.333 * GH - 0.333 * GK - 0.333 * GL + 0.333 * GM + 0.333 * HK - 0.333 * HL
- 0.333 * HM - 0.333 * KL + 0.333 * KM - 0.333 * LM

[AK] = AK - 0.333 * BC - 0.333 * BD + 0.333 * BE - 0.333 * BF - 0.333 * BG
- 0.333 * BH + 0.333 * BJ - 0.333 * BL - 0.333 * BM + 0.333 * CD + 0.333 * CE
+ 0.333 * CF + 0.333 * CG - 0.333 * CH + 0.333 * CJ - 0.333 * CL + 0.333 * CM
+ 0.333 * DE + 0.333 * DF - 0.333 * DG + 0.333 * DH + 0.333 * DJ + 0.333 * DL
- 0.333 * DM - 0.333 * EF + 0.333 * EG - 0.333 * EH + 0.333 * EJ + 0.333 * EL
- 0.333 * EM - 0.333 * FG - 0.333 * FH - 0.333 * FJ - 0.333 * FL - 0.333 * FM
- 0.333 * GH - 0.333 * GJ + 0.333 * GL + 0.333 * GM + 0.333 * HJ + 0.333 * HL
+ 0.333 * HM - 0.333 * JL + 0.333 * JM - 0.333 * LM

[AL] = AL - 0.333 * BC + 0.333 * BD + 0.333 * BE + 0.333 * BF - 0.333 * BG
+ 0.333 * BH + 0.333 * BJ - 0.333 * BK + 0.333 * BM - 0.333 * CD + 0.333 * CE
+ 0.333 * CF - 0.333 * CG - 0.333 * CH - 0.333 * CJ - 0.333 * CK - 0.333 * CM
+ 0.333 * DE + 0.333 * DF + 0.333 * DG - 0.333 * DH + 0.333 * DJ + 0.333 * DK
- 0.333 * DM + 0.333 * EF - 0.333 * EG + 0.333 * EH - 0.333 * EJ + 0.333 * EK
- 0.333 * EM + 0.333 * FG - 0.333 * FH - 0.333 * FJ - 0.333 * FK + 0.333 * FM
- 0.333 * GH - 0.333 * GJ + 0.333 * GK + 0.333 * GM - 0.333 * HJ + 0.333 * HK
+ 0.333 * HM - 0.333 * JK - 0.333 * JM - 0.333 * KM

[AM] = AM + 0.333 * BC - 0.333 * BD + 0.333 * BE + 0.333 * BF + 0.333 * BG
- 0.333 * BH + 0.333 * BJ - 0.333 * BK + 0.333 * BL - 0.333 * CD + 0.333 * CE
+ 0.333 * CF - 0.333 * CG + 0.333 * CH + 0.333 * CJ + 0.333 * CK - 0.333 * CL
- 0.333 * DE + 0.333 * DF - 0.333 * DG - 0.333 * DH + 0.333 * DJ - 0.333 * DK
- 0.333 * DL - 0.333 * EF - 0.333 * EG - 0.333 * EH - 0.333 * EJ - 0.333 * EK
- 0.333 * EL - 0.333 * FG + 0.333 * FH + 0.333 * FJ - 0.333 * FK + 0.333 * FL
- 0.333 * GH + 0.333 * GJ + 0.333 * GK + 0.333 * GL - 0.333 * HJ + 0.333 * HK
+ 0.333 * HL + 0.333 * JK - 0.333 * JL - 0.333 * KL

Forward selection regression should be used for analyzing the experiment. Hierarchy should be considered.

9.31. Suppose that you must design an experiment to investigate nine continuous factors. It is thought that running all factors at two levels is adequate but that all two-factor interactions are of interest.

(a) How many runs are required to estimate all main effects and two factor-interactions?

1 for the intercept term, 9 for the main effects, and 36 ($\frac{9!}{2!(9-2)!}$) for the two-factor interactions, for a total of 46 runs.

(b) Find a minimum-run D-optimal design that is suitable for this problem.

A	B	C	D	E	F	G	H	J
+	-	-	+	+	-	-	-	-
+	-	+	-	+	-	+	+	+
+	+	-	+	+	-	-	+	+
+	+	+	-	+	+	+	-	+
+	+	-	-	+	-	-	-	+
-	-	-	-	+	+	-	+	+
+	-	+	+	+	+	-	+	+
+	+	+	+	+	+	-	-	-
+	-	-	+	-	-	+	-	+
+	-	-	+	-	+	-	+	-
-	+	-	-	+	+	-	-	-
-	-	+	-	-	-	-	+	-
-	+	+	-	+	-	-	+	+
+	+	+	+	-	+	+	-	+
-	-	-	+	-	+	+	+	+
+	+	-	-	-	+	+	-	-
-	+	-	+	-	+	+	+	-
+	+	-	+	+	+	+	+	-
+	-	+	-	-	-	-	-	-
-	-	+	-	-	+	+	-	+
-	-	+	-	+	+	-	-	-
+	-	-	+	-	+	-	-	+
+	+	-	-	-	+	-	+	+
-	-	-	-	+	+	+	+	-
-	+	-	-	-	+	+	-	+
-	-	-	+	+	+	+	-	+
+	-	-	-	-	-	-	+	+
-	-	+	+	+	+	+	+	+
-	+	+	-	-	-	-	-	-
-	-	+	+	-	+	+	-	-
+	+	+	-	-	-	+	-	+
-	+	+	+	-	-	+	+	+
-	+	+	+	+	-	-	+	-
+	-	-	-	-	-	+	+	-
+	-	+	+	+	+	+	-	-
-	+	+	+	-	+	-	+	+
-	+	-	-	+	-	+	+	+
-	+	-	+	-	-	-	-	+
+	+	+	+	+	-	+	-	-
-	-	-	+	+	-	+	+	-
-	-	+	-	+	-	+	-	-
+	+	+	+	-	-	-	+	-
-	-	-	-	-	-	-	-	-
-	-	+	+	+	-	-	-	+
+	+	+	-	-	+	+	+	-
+	+	-	-	+	-	-	+	-

9.32. Suppose that you must design an experiment to investigate seven continuous factors. Running all factors at two levels is thought to be appropriate but that only the two-factor interactions involving factor A are of interest.

(a) How many runs are required to estimate all main effects and two factor-interactions?

1 for the intercept term, 7 for the main effects, and 6 for the two-factor interactions involving A, for a total of 14 runs.

(b) Find a minimum-run D-optimal design that is suitable for this problem.

A	B	C	D	E	F	G
-	-	-	+	-	-	+
+	+	-	+	-	-	-
-	+	+	-	-	-	+
-	+	-	-	+	+	+
+	+	+	+	+	+	-
+	-	+	+	-	-	+
-	+	+	+	+	-	-
+	-	-	-	+	+	+
+	+	+	-	-	+	+
+	+	+	-	+	-	-
-	+	-	-	-	-	-
+	-	+	-	-	+	-
-	-	+	-	+	+	-
-	+	+	+	-	+	-

9.33. Suppose that you must design an experiment to investigate six continuous factors. It is thought that running all factors at two levels is adequate but that only the AB, AC, and AD two-factor interactions are of interest.

(a) How many runs are required to estimate all main effects and two factor-interactions?

1 for the intercept term, 6 for the main effects, and 3 for the two-factor interactions involving AB, AC and AD, for a total of 10 runs.

(b) Find a minimum-run D-optimal design that is suitable for this problem.

A	B	C	D	E	F
+	+	-	-	-	-
+	+	+	+	+	+
-	-	-	+	-	+
+	-	+	+	-	-
+	-	-	+	+	+
-	-	+	+	+	-
-	-	+	-	-	+
-	+	-	-	+	-
+	-	+	-	+	+
-	+	+	+	-	+

9.34. Suppose that you must design an experiment with six categorical factors. Factor A has six levels, factor B has five levels, factor C has five levels, factor D has three levels, and factors E and F have two levels. You are interested in main effects and two factor interactions.

(a) How many runs are required to estimate all of the effects that are of interest?

Intercept	1	*AB*	20	*BE*	4
A	5	*AC*	20	*BF*	4
B	4	*AD*	10	*CD*	8
C	4	*AE*	5	*CE*	4
D	2	*AF*	5	*CF*	4
E	1	*BC*	16	*DE*	2
F	1	*BD*	8	*DF*	2
				EF	1
Total	131				

(b) Find a D-optimal design suitable for this problem.

Order	A	B	C	D	E	F
1	L2	L5	L4	L1	L1	L1
2	L5	L3	L5	L2	L1	L2
3	L3	L2	L4	L3	L2	L1
4	L5	L4	L2	L1	L1	L2
5	L1	L2	L3	L2	L2	L2
6	L4	L1	L3	L1	L1	L1
7	L1	L3	L2	L2	L1	L2
8	L6	L5	L1	L1	L2	L1
9	L5	L4	L5	L1	L2	L2
10	L4	L2	L5	L2	L2	L1
11	L2	L2	L1	L3	L1	L1
12	L5	L1	L1	L2	L2	L2
13	L2	L3	L5	L2	L2	L2
14	L4	L5	L3	L2	L2	L2
15	L3	L4	L2	L3	L2	L2
16	L6	L5	L3	L1	L1	L2
17	L5	L2	L1	L2	L2	L2
18	L4	L1	L4	L3	L1	L1
19	L5	L3	L2	L3	L2	L2
20	L1	L4	L5	L2	L1	L1
21	L6	L4	L4	L1	L1	L1
22	L4	L2	L3	L1	L2	L2
23	L2	L3	L1	L2	L2	L1
24	L3	L5	L5	L3	L1	L2
25	L2	L1	L3	L2	L1	L2
26	L2	L1	L5	L1	L2	L1
27	L6	L2	L2	L2	L2	L2
28	L1	L4	L1	L1	L2	L2
29	L3	L5	L4	L2	L2	L1
30	L4	L3	L3	L3	L2	L2
31	L4	L1	L5	L2	L1	L2
32	L2	L3	L2	L1	L1	L1
33	L6	L2	L4	L3	L1	L2
34	L5	L5	L1	L3	L2	L1
35	L1	L5	L2	L1	L1	L1
36	L3	L2	L5	L3	L2	L2
37	L4	L4	L2	L2	L2	L1
38	L2	L5	L3	L3	L2	L2
39	L1	L2	L4	L1	L1	L2
40	L6	L2	L1	L2	L1	L2
41	L3	L2	L1	L1	L1	L2
42	L5	L5	L3	L3	L1	L1
43	L1	L4	L4	L3	L2	L2
44	L5	L1	L2	L2	L1	L1
45	L3	L3	L3	L1	L2	L1
46	L6	L4	L5	L1	L1	L2
47	L2	L2	L3	L1	L1	L2
48	L4	L1	L2	L1	L2	L2

Order	A	B	C	D	E	F
67	L2	L1	L4	L1	L2	L2
68	L6	L3	L5	L3	L2	L1
69	L3	L2	L2	L2	L1	L2
70	L5	L4	L3	L2	L2	L1
71	L4	L4	L5	L3	L1	L2
72	L1	L1	L1	L1	L1	L1
73	L4	L3	L4	L2	L1	L2
74	L6	L5	L2	L3	L2	L1
75	L3	L2	L3	L2	L1	L1
76	L5	L5	L2	L1	L2	L2
77	L6	L1	L3	L3	L2	L1
78	L2	L4	L1	L3	L1	L2
79	L6	L2	L5	L1	L2	L1
80	L1	L3	L4	L2	L2	L1
81	L1	L5	L1	L3	L1	L1
82	L2	L3	L4	L3	L2	L1
83	L3	L4	L5	L1	L1	L1
84	L4	L4	L3	L2	L1	L2
85	L2	L1	L2	L1	L1	L2
86	L3	L3	L2	L2	L2	L1
87	L4	L1	L1	L3	L1	L2
88	L5	L1	L4	L1	L2	L1
89	L2	L4	L5	L3	L2	L1
90	L6	L5	L5	L2	L1	L1
91	L4	L4	L1	L2	L2	L2
92	L2	L5	L2	L2	L1	L1
93	L1	L5	L3	L1	L2	L1
94	L5	L1	L5	L3	L1	L1
95	L3	L3	L4	L3	L1	L2
96	L6	L1	L4	L2	L2	L1
97	L1	L3	L3	L3	L1	L1
98	L1	L5	L5	L2	L2	L2
99	L6	L4	L1	L3	L2	L1
100	L4	L2	L2	L3	L1	L2
101	L2	L2	L2	L1	L2	L1
102	L3	L4	L4	L2	L1	L2
103	L6	L2	L3	L3	L1	L1
104	L4	L1	L1	L2	L2	L1
105	L6	L5	L5	L3	L2	L2
106	L4	L2	L1	L1	L1	L1
107	L2	L1	L2	L3	L2	L1
108	L3	L3	L5	L2	L1	L1
109	L4	L4	L4	L1	L2	L1
110	L5	L2	L3	L3	L1	L2
111	L3	L1	L3	L2	L2	L2
112	L2	L2	L5	L3	L1	L2
113	L6	L3	L4	L1	L2	L2
114	L1	L4	L2	L3	L1	L1

49	L5	L3	L1	L1	L1	L2
50	L2	L4	L4	L2	L1	L1
51	L4	L5	L4	L3	L1	L2
52	L6	L3	L3	L2	L1	L1
53	L3	L1	L1	L3	L1	L1
54	L1	L2	L5	L1	L2	L1
55	L2	L4	L2	L2	L2	L2
56	L5	L2	L2	L1	L1	L1
57	L4	L3	L1	L3	L1	L1
58	L6	L4	L3	L2	L2	L2
59	L1	L1	L5	L3	L2	L2
60	L3	L5	L4	L1	L2	L2
61	L4	L5	L5	L1	L1	L1
62	L3	L3	L1	L3	L2	L2
63	L2	L2	L4	L2	L1	L2
64	L3	L4	L3	L3	L1	L1
65	L6	L1	L1	L1	L2	L2
66	L5	L4	L1	L2	L1	L1

115	L3	L4	L1	L2	L2	L1
116	L5	L5	L4	L2	L2	L2
117	L2	L4	L3	L1	L2	L1
118	L1	L2	L1	L3	L2	L1
119	L4	L3	L5	L1	L2	L1
120	L6	L1	L2	L3	L1	L2
121	L5	L2	L5	L3	L2	L1
122	L5	L3	L3	L1	L2	L1
123	L1	L5	L4	L2	L1	L2
124	L6	L4	L2	L1	L2	L1
125	L2	L5	L1	L1	L2	L2
126	L3	L5	L1	L2	L1	L2
127	L4	L3	L2	L1	L1	L2
128	L1	L1	L2	L2	L2	L1
129	L5	L4	L4	L3	L1	L1
130	L2	L2	L3	L2	L2	L1
131	L3	L3	L5	L1	L1	L2

(c) Suppose that the experimenter decides that this is too many runs. What strategy would you recommend?

A design for main effects could be initially run with only 18 runs which would identify the significant factors. This could be followed by a second experiment that would include the significant factors and levels of interest.

Chapter 10
Fitting Regression Models
Solutions

10.1. The tensile strength of a paper product is related to the amount of hardwood in the pulp. Ten samples are produced in the pilot plant, and the data obtained are shown in the following table.

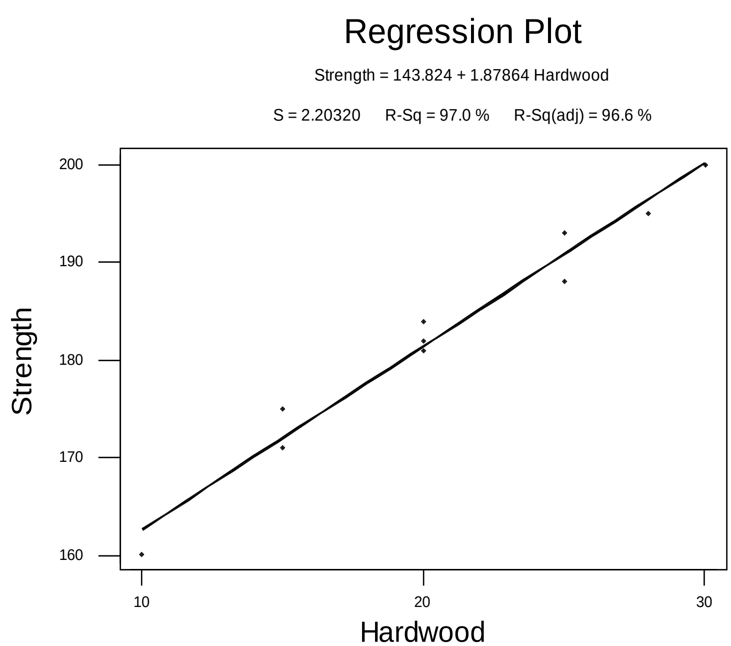
Strength	Percent Hardwood	Strength	Percent Hardwood
160	10	181	20
171	15	188	25
175	15	193	25
182	20	195	28
184	20	200	30

(a) Fit a linear regression model relating strength to percent hardwood.

Minitab Output

Regression Analysis: Strength versus Hardwood

```
The regression equation is
Strength = 144 + 1.88 Hardwood

Predictor        Coef     SE Coef          T        P
Constant      143.824       2.522      57.04    0.000
Hardwood       1.8786      0.1165      16.12    0.000

S = 2.203              R-Sq = 97.0%         R-Sq(adj) = 96.6%
PRESS = 66.2665        R-Sq(pred) = 94.91%
```



(b) Test the model in part (a) for significance of regression.

Minitab Output

```
Analysis of Variance

Source            DF        SS        MS        F        P
Regression         1    1262.1    1262.1   260.00    0.000
Residual Error     8      38.8       4.9
  Lack of Fit      4      13.7       3.4     0.54    0.716
  Pure Error       4      25.2       6.3
Total              9    1300.9

3 rows with no replicates

No evidence of lack of fit (P > 0.1)
```

(c) Find a 95 percent confidence interval on the parameter β_1.

The 95 percent confidence interval is:

$$\hat{\beta}_1 - t_{\alpha/2,n-p}se\left(\hat{\beta}_1\right) \leq \beta_1 \leq \hat{\beta}_1 + t_{\alpha/2,n-p}se\left(\hat{\beta}_1\right)$$
$$1.8786 - 2.3060(0.1165) \leq \beta_1 \leq 1.8786 + 2.3060(0.1165)$$
$$1.6900 \leq \beta_1 \leq 2.1473$$

10.2. A plant distills liquid air to produce oxygen, nitrogen, and argon. The percentage of impurity in the oxygen is thought to be linearly related to the amount of impurities in the air as measured by the "pollution count" in part per million (ppm). A sample of plant operating data is shown below.

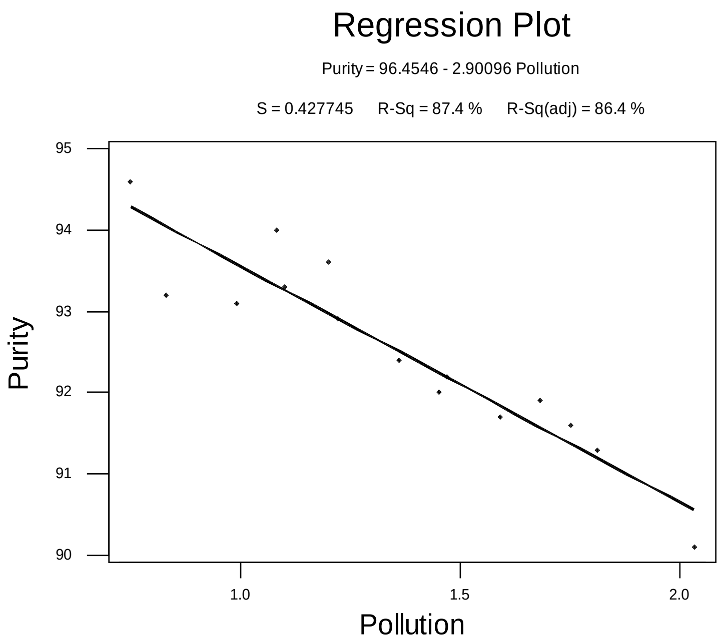
Purity(%)	93.3	92.0	92.4	91.7	94.0	94.6	93.6	93.1	93.2	92.9	92.2	91.3	90.1	91.6	91.9
Pollution count (ppm)	1.10	1.45	1.36	1.59	1.08	0.75	1.20	0.99	0.83	1.22	1.47	1.81	2.03	1.75	1.68

(a) Fit a linear regression model to the data.

Minitab Output

Regression Analysis: Purity versus Pollution

```
The regression equation is
Purity = 96.5 - 2.90 Pollution

Predictor        Coef     SE Coef          T        P
Constant      96.4546      0.4282     225.24    0.000
Pollutio      -2.9010      0.3056      -9.49    0.000

S = 0.4277            R-Sq = 87.4%        R-Sq(adj) = 86.4%
PRESS = 3.43946       R-Sq(pred) = 81.77%
```



(b) Test for significance of regression.

Minitab Output

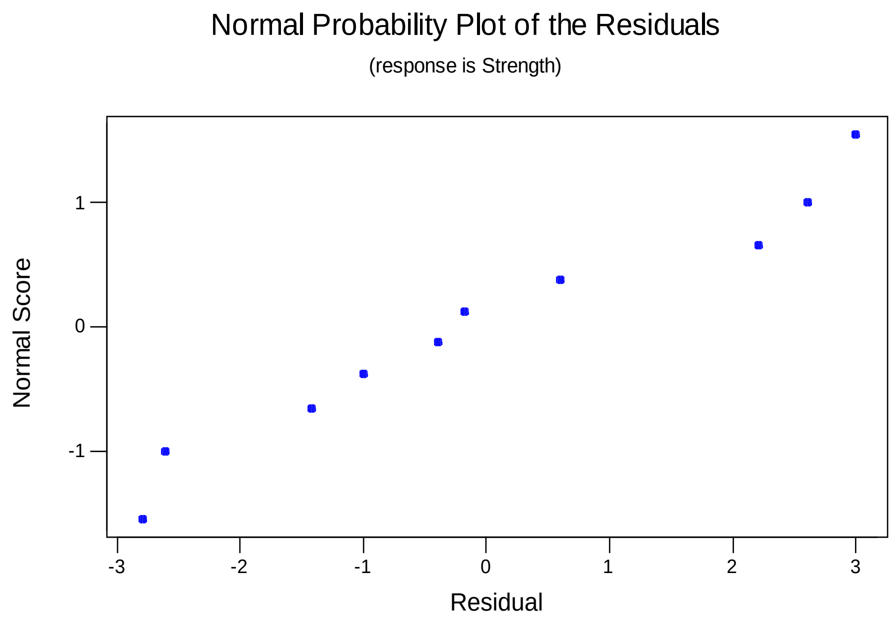
```
Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      16.491      16.491     90.13    0.000
Residual Error    13       2.379       0.183
Total             14      18.869

No replicates. Cannot do pure error test.

No evidence of lack of fit (P > 0.1)
```

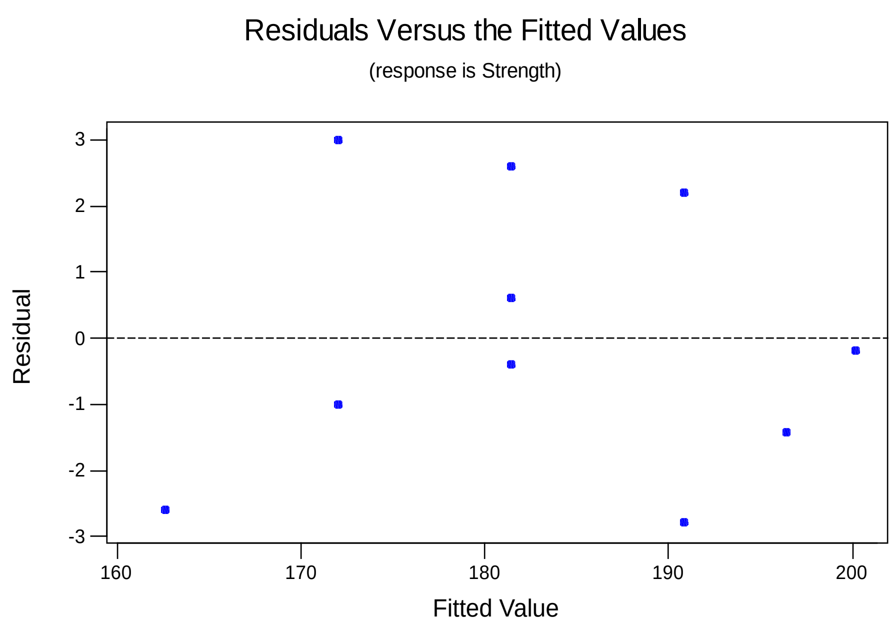
(c) Find a 95 percent confidence interval on β_1.

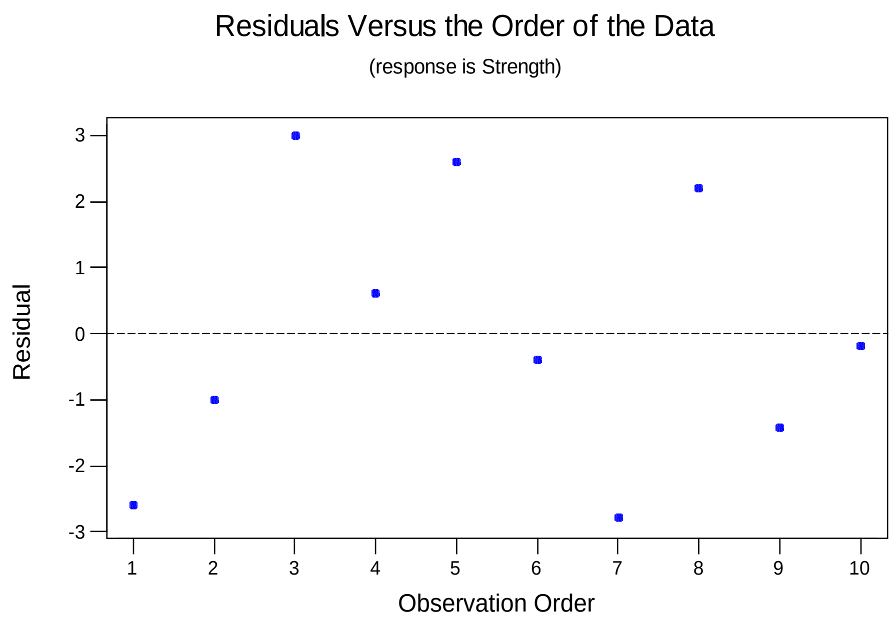
The 95 percent confidence interval is:

$$\hat{\beta}_1 - t_{\alpha/2,n-p} se(\hat{\beta}_1) \le \beta_1 \le \hat{\beta}_1 + t_{\alpha/2,n-p} se(\hat{\beta}_1)$$
$$-2.9010 - 2.1604(0.3056) \le \beta_1 \le -2.9010 + 2.1604(0.3056)$$
$$-3.5612 \le \beta_1 \le -2.2408$$

10.3. Plot the residuals from Problem 10.1 and comment on model adequacy.

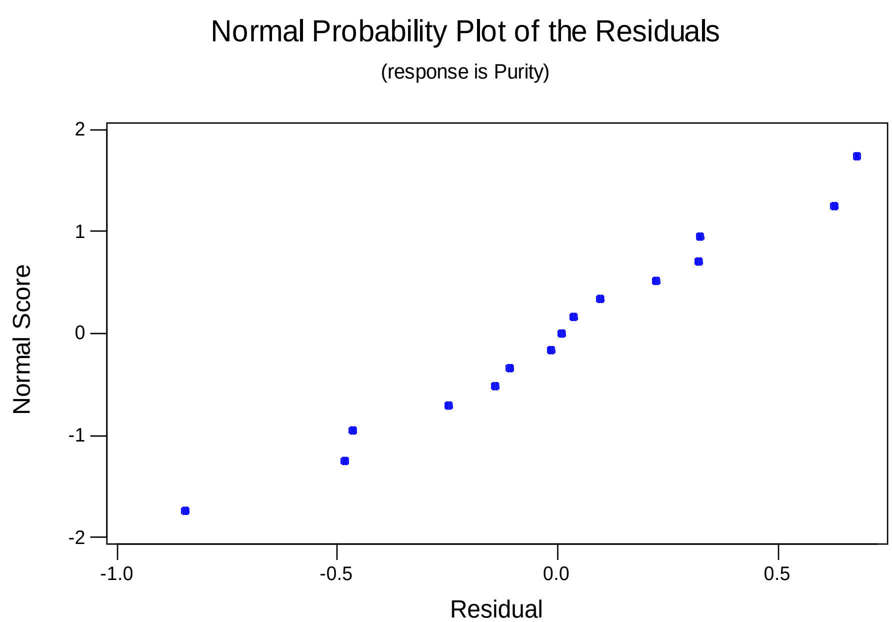


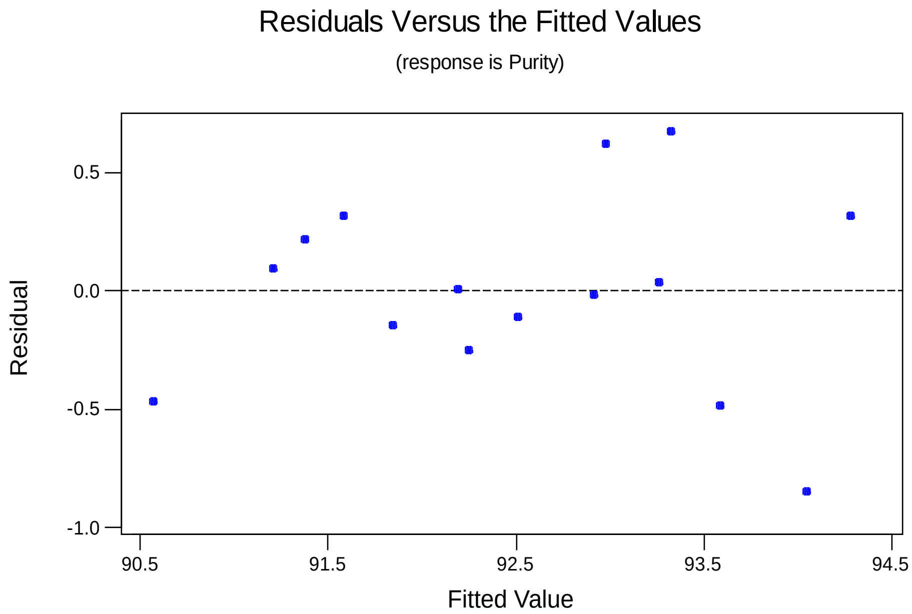


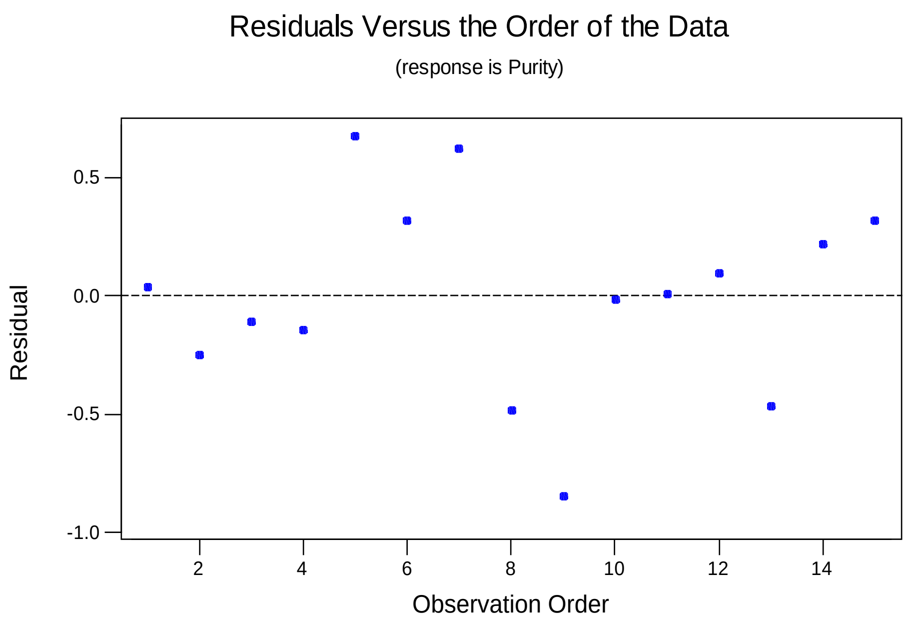


There is nothing unusual about the residual plots. The underlying assumptions have been met.

10.4. Plot the residuals from Problem 10.2 and comment on model adequacy.







There is nothing unusual about the residual plots. The underlying assumptions have been met.

10.5. Using the results of Problem 10.1, test the regression model for lack of fit.

The Minitab output below identifies no evidence of lack of fit.

Minitab Output

```
Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      1262.1      1262.1    260.00    0.000
Residual Error     8        38.8         4.9
  Lack of Fit      4        13.7         3.4      0.54    0.716
  Pure Error       4        25.2         6.3
Total              9      1300.9

3 rows with no replicates

No evidence of lack of fit (P > 0.1)
```

10.6. A study was performed on wear of a bearing y and its relationship to x_1 = oil viscosity and x_2 = load. The following data were obtained.

y	x_1	x_2
193	1.6	851
230	15.5	816
172	22.0	1058
91	43.0	1201
113	33.0	1357
125	40.0	1115

(a) Fit a multiple linear regression model to the data.

Minitab Output

Regression Analysis: Wear versus Viscosity, Load

```
The regression equation is
Wear = 351 - 1.27 Viscosity - 0.154 Load

Predictor        Coef     SE Coef          T        P       VIF
Constant       350.99       74.75       4.70    0.018
Viscosit       -1.272       1.169      -1.09    0.356       2.6
Load         -0.15390     0.08953      -1.72    0.184       2.6

S = 25.50               R-Sq = 86.2%         R-Sq(adj) = 77.0%
PRESS = 12696.7         R-Sq(pred) = 10.03%
```

(b) Test for significance of regression.

Minitab Output

```
Analysis of Variance

Source            DF          SS          MS         F        P
Regression         2     12161.6      6080.8      9.35    0.051
Residual Error     3      1950.4       650.1
Total              5     14112.0

No replicates. Cannot do pure error test.

Source       DF      Seq SS
Viscosit      1     10240.4
Load          1      1921.2

* Not enough data for lack of fit test
```

(c) Compute *t* statistics for each model parameter. What conclusions can you draw?

Minitab Output

```
Regression Analysis: Wear versus Viscosity, Load

The regression equation is
Wear = 351 - 1.27 Viscosity - 0.154 Load

Predictor        Coef     SE Coef          T        P       VIF
Constant       350.99       74.75       4.70    0.018
Viscosit       -1.272       1.169      -1.09    0.356       2.6
Load         -0.15390     0.08953      -1.72    0.184       2.6

S = 25.50              R-Sq = 86.2%         R-Sq(adj) = 77.0%
PRESS = 12696.7        R-Sq(pred) = 10.03%
```

The *t*-tests are shown in part (a). Notice that overall regression is significant (part(b)), but neither variable has a large *t*-statistic. This could be an indicator that the regressors are nearly linearly dependent.

10.7. The brake horsepower developed by an automobile engine on a dynamometer is thought to be a function of the engine speed in revolutions per minute (rpm), the road octane number of the fuel, and the engine compression. An experiment is run in the laboratory and the data that follow are collected.

Brake Horsepower	rpm	Road Octane Number	Compression
225	2000	90	100
212	1800	94	95
229	2400	88	110
222	1900	91	96
219	1600	86	100
278	2500	96	110
246	3000	94	98
237	3200	90	100
233	2800	88	105
224	3400	86	97
223	1800	90	100
230	2500	89	104

(a) Fit a multiple linear regression model to the data.

Minitab Output

```
Regression Analysis: Horsepower versus rpm, Octane, Compression

The regression equation is
Horsepower = - 266 + 0.0107 rpm + 3.13 Octane + 1.87 Compression

Predictor        Coef     SE Coef          T        P       VIF
Constant      -266.03       92.67      -2.87    0.021
rpm          0.010713    0.004483       2.39    0.044       1.0
Octane         3.1348      0.8444       3.71    0.006       1.0
Compress       1.8674      0.5345       3.49    0.008       1.0

S = 8.812              R-Sq = 80.7%         R-Sq(adj) = 73.4%
PRESS = 2494.05        R-Sq(pred) = 22.33%
```

(b) Test for significance of regression. What conclusions can you draw?

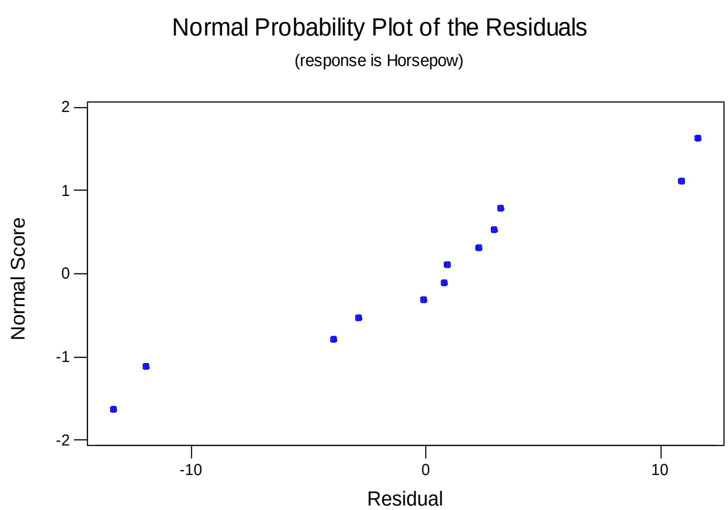
Minitab Output

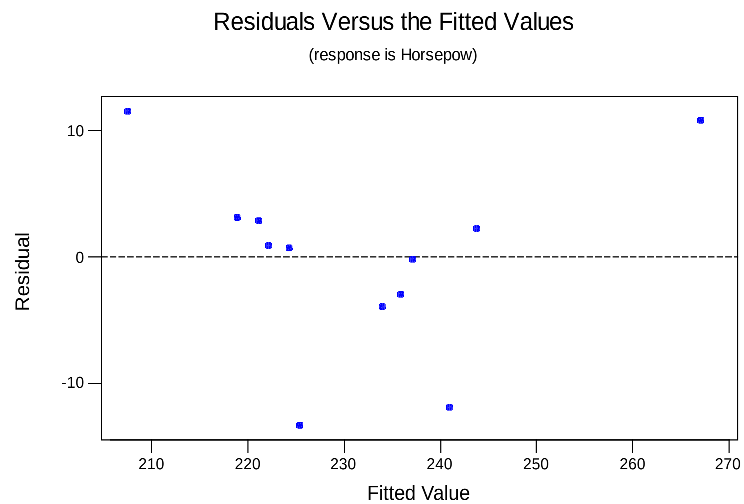
```
Analysis of Variance

Source           DF          SS          MS         F        P
Regression        3     2589.73      863.24     11.12    0.003
Residual Error    8      621.27       77.66
Total            11     3211.00
r No replicates. Cannot do pure error test.

Source       DF      Seq SS
rpm           1      509.35
Octane        1     1132.56
Compress      1      947.83

Lack of fit test
Possible interactions with variable Octane (P-Value = 0.028)
Possible lack of fit at outer X-values       (P-Value = 0.000)
Overall lack of fit test is significant at P = 0.000
```

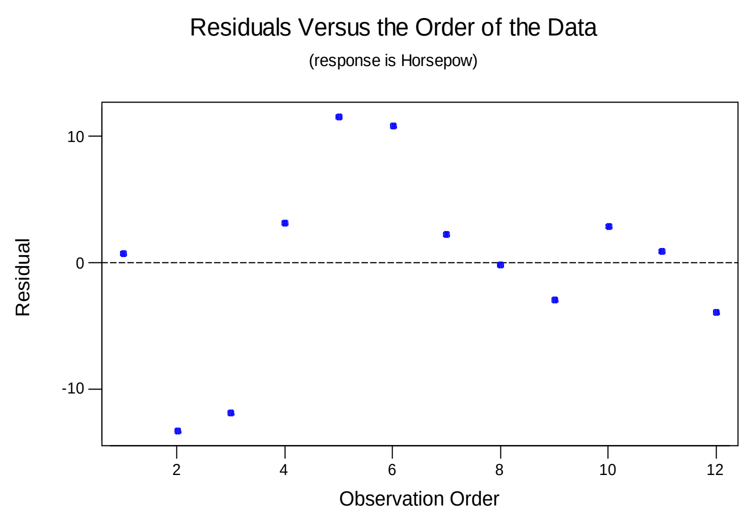
(c) Based on *t* tests, do you need all three regressor variables in the model?

Yes, all of the regressor variables are important.

10.8. Analyze the residuals from the regression model in Problem 10.7. Comment on model adequacy.







The normal probability plot is satisfactory, as is the plot of residuals versus run order (assuming that observation order is run order). The plot of residuals versus predicted exhibits a slight "bow" shape. This could be an indication of lack of fit. It might be useful to add some interaction terms to the model.

10.9. The yield of a chemical process is related to the concentration of the reactant and the operating temperature. An experiment has been conducted with the following results.

Yield	Concentration	Temperature
81	1.00	150
89	1.00	180
83	2.00	150
91	2.00	180
79	1.00	150
87	1.00	180
84	2.00	150
90	2.00	180

(a) Suppose we wish to fit a main effects model to this data. Set up the **X'X** matrix using the data exactly as it appears in the table.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1.00 & 1.00 & 2.00 & 2.00 & 1.00 & 1.00 & 2.00 & 2.00 \\ 150 & 180 & 150 & 180 & 150 & 180 & 150 & 180 \end{bmatrix} \begin{bmatrix} 1 & 1.00 & 150 \\ 1 & 1.00 & 180 \\ 1 & 2.00 & 150 \\ 1 & 2.00 & 180 \\ 1 & 1.00 & 150 \\ 1 & 1.00 & 180 \\ 1 & 2.00 & 150 \\ 1 & 2.00 & 180 \end{bmatrix} = \begin{bmatrix} 8 & 12 & 1320 \\ 12 & 20 & 1980 \\ 1320 & 1980 & 219600 \end{bmatrix}$$

(b) Is the matrix you obtained in part (a) diagonal? Discuss your response.

The **X'X** is not diagonal, even though an orthogonal design has been used. The reason is that we have worked with the natural factor levels, not the orthogonally coded variables.

(c) Suppose we write our model in terms of the "usual" coded variables

$$x_1 = \frac{Conc - 1.5}{0.5}, \ x_2 = \frac{Temp - 165}{15}$$

Set up the **X'X** matrix for the model in terms of these coded variables. Is this matrix diagonal? Discuss your response.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

The **X'X** matrix is diagonal because we have used the orthogonally coded variables.

(d) Define a new set of coded variables

$$x_1 = \frac{Conc - 1.0}{1.0}, \ x_2 = \frac{Temp - 150}{30}$$

Set up the **X'X** matrix for the model in terms of this set of coded variables. Is this matrix diagonal? Discuss your response.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 8 & 4 & 4 \\ 4 & 4 & 2 \\ 4 & 2 & 4 \end{bmatrix}$$

The **X'X** is not diagonal, even though an orthogonal design has been used. The reason is that we have not used orthogonally coded variables.

(e) Summarize what you have learned from this problem about coding the variables.

If the design is orthogonal, use the orthogonal coding. This not only makes the analysis somewhat easier, but it also results in model coefficients that are easier to interpret because they are both dimensionless and uncorrelated.

10.10. Consider the 2^4 factorial experiment in Example 6-2. Suppose that the last observation in missing. Reanalyze the data and draw conclusions. How do these conclusions compare with those from the original example?

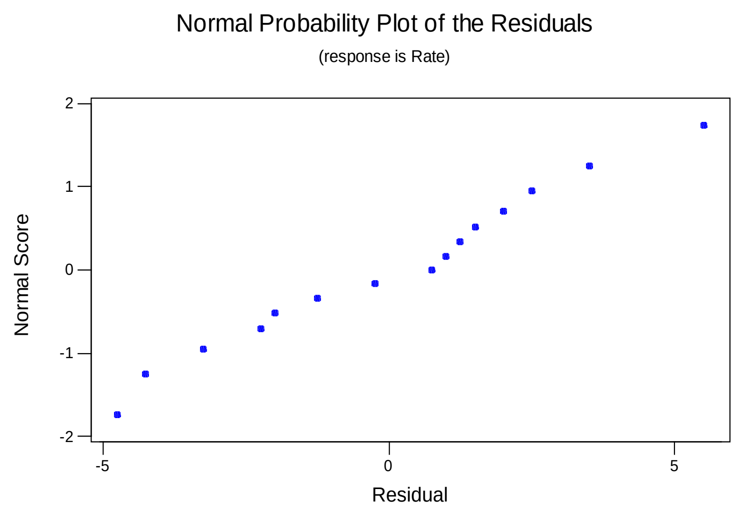
The regression analysis with the one data point missing indicates that the same effects are important.

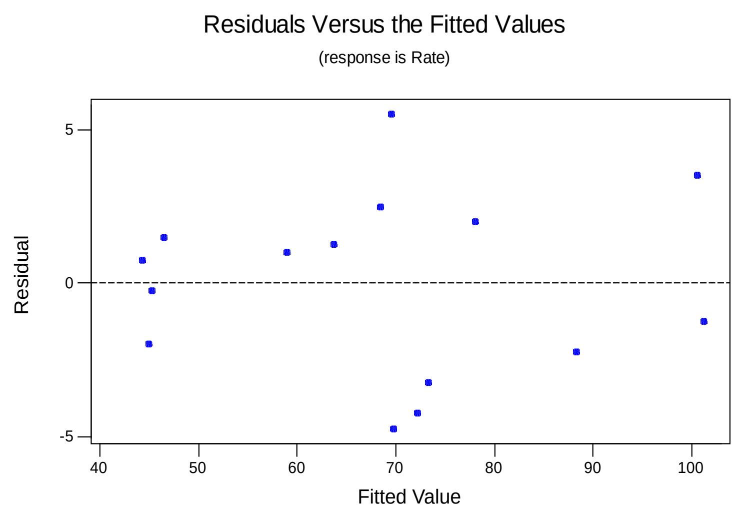
Minitab Output

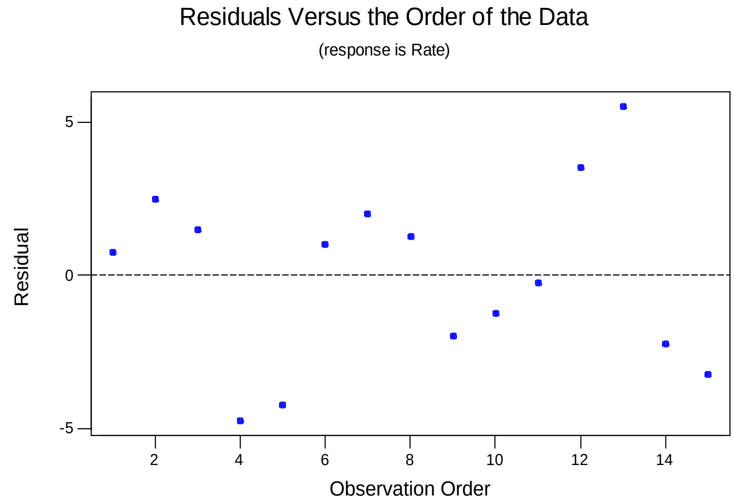
Regression Analysis: Rate versus A, B, C, D, AB, AC, AD, BC, BD, CD

```
The regression equation is
Rate = 69.8 + 10.5 A + 1.25 B + 4.63 C + 7.00 D - 0.25 AB - 9.38 AC + 8.00 AD
           + 0.87 BC - 0.50 BD - 0.87 CD

Predictor        Coef     SE Coef          T        P       VIF
Constant       69.750       1.500      46.50    0.000
A              10.500       1.500       7.00    0.002       1.1
B               1.250       1.500       0.83    0.452       1.1
C               4.625       1.500       3.08    0.037       1.1
D               7.000       1.500       4.67    0.010       1.1
AB             -0.250       1.500      -0.17    0.876       1.1
AC             -9.375       1.500      -6.25    0.003       1.1
AD              8.000       1.500       5.33    0.006       1.1
BC              0.875       1.500       0.58    0.591       1.1
BD             -0.500       1.500      -0.33    0.756       1.1
CD             -0.875       1.500      -0.58    0.591       1.1

S = 5.477              R-Sq = 97.6%         R-Sq(adj) = 91.6%
PRESS = 1750.00        R-Sq(pred) = 65.09%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression        10     4893.33      489.33     16.31    0.008
Residual Error     4      120.00       30.00
Total             14     5013.33

No replicates. Cannot do pure error test.

Source       DF      Seq SS
A             1     1414.40
B             1        4.01
C             1      262.86
```

D	1	758.88
AB	1	0.06
AC	1	1500.63
AD	1	924.50
BC	1	16.07
BD	1	1.72
CD	1	10.21







The residual plots are acceptable; therefore, the underlying assumptions are valid.

10.11. Consider the 2^4 factorial experiment in Example 6-2. Suppose that the last two observations are missing. Reanalyze the data and draw conclusions. How do these conclusions compare with those from the original example?

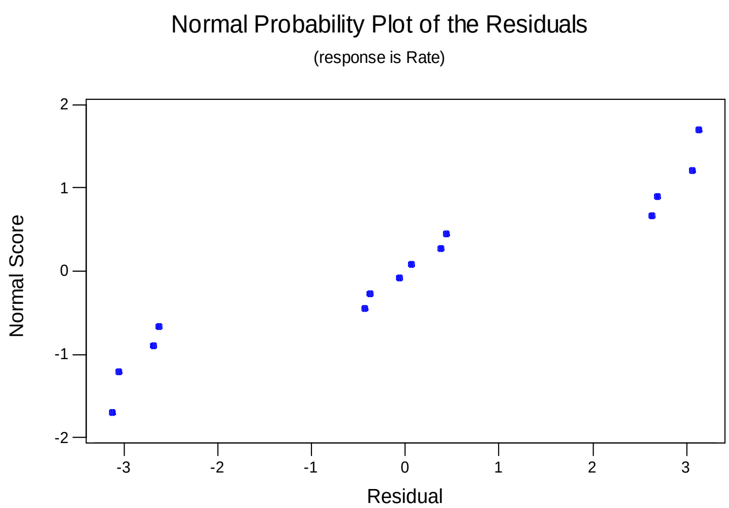
The regression analysis with two data points missing indicates that the same effects are important.

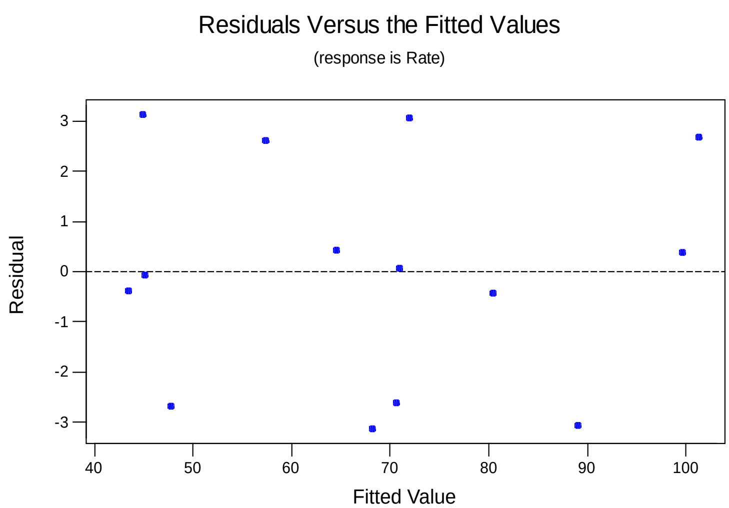
Minitab Output

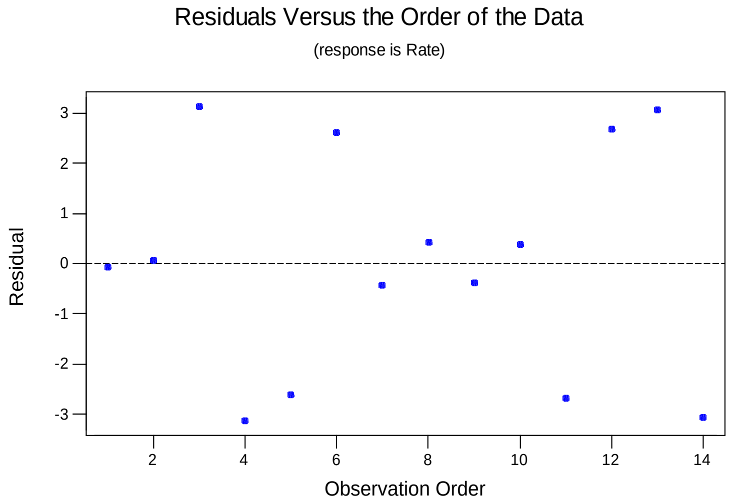
Regression Analysis: Rate versus A, B, C, D, AB, AC, AD, BC, BD, CD

```
The regression equation is
Rate = 71.4 + 10.1 A + 2.87 B + 6.25 C + 8.62 D - 0.66 AB - 9.78 AC + 7.59 AD
           + 2.50 BC + 1.12 BD + 0.75 CD

Predictor        Coef     SE Coef         T        P       VIF
Constant       71.375       1.673     42.66    0.000
A              10.094       1.323      7.63    0.005       1.1
B               2.875       1.673      1.72    0.184       1.7
C               6.250       1.673      3.74    0.033       1.7
D               8.625       1.673      5.15    0.014       1.7
AB             -0.656       1.323     -0.50    0.654       1.1
AC             -9.781       1.323     -7.39    0.005       1.1
AD              7.594       1.323      5.74    0.010       1.1
BC              2.500       1.673      1.49    0.232       1.7
BD              1.125       1.673      0.67    0.549       1.7
CD              0.750       1.673      0.45    0.684       1.7

S = 4.732               R-Sq = 98.7%         R-Sq(adj) = 94.2%
PRESS = 1493.06         R-Sq(pred) = 70.20%

Analysis of Variance

Source            DF          SS         MS         F         P
Regression        10     4943.17     494.32     22.07     0.014
Residual Error     3       67.19      22.40
Total             13     5010.36

No replicates. Cannot do pure error test.
```

Source	DF	Seq SS
A	1	1543.50
B	1	1.52
C	1	177.63
D	1	726.01
AB	1	1.17
AC	1	1702.53
AD	1	738.11
BC	1	42.19
BD	1	6.00
CD	1	4.50







The residual plots are acceptable; therefore, the underlying assumptions are valid.

10.12. Given the following data, fit the second-order polynomial regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \varepsilon$$

y	x_1	x_2
26	1.0	1.0
24	1.0	1.0
175	1.5	4.0
160	1.5	4.0
163	1.5	4.0
55	0.5	2.0
62	1.5	2.0
100	0.5	3.0
26	1.0	1.5
30	0.5	1.5
70	1.0	2.5
71	0.5	2.5

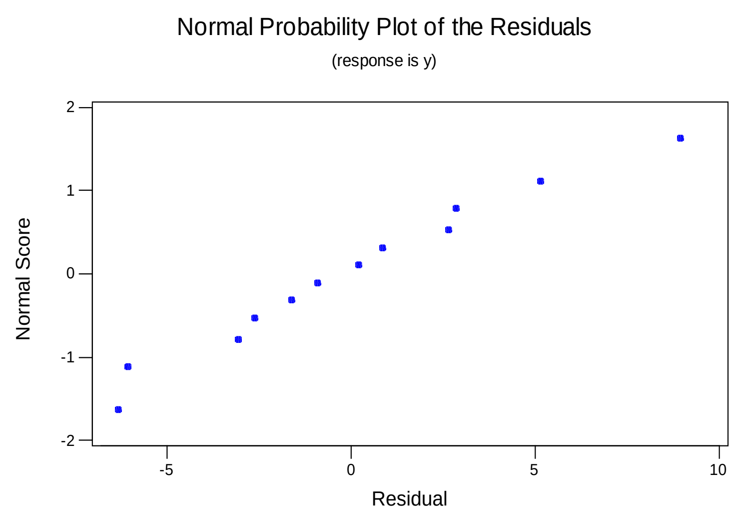
After you have fit the model, test for significance of regression.

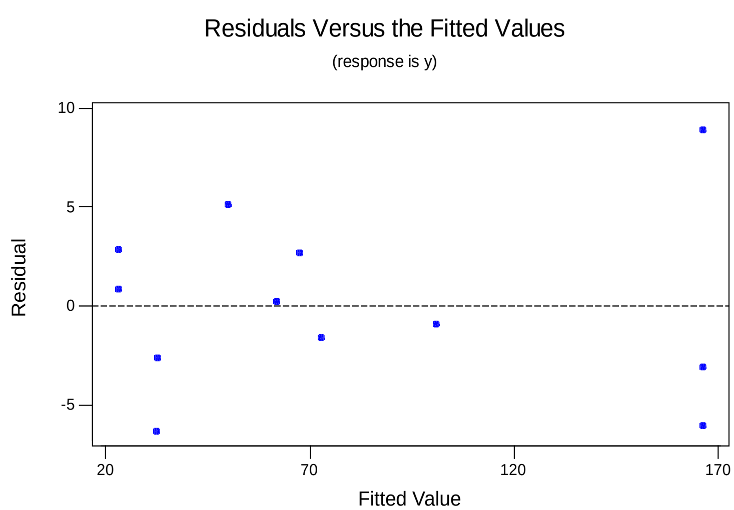
Minitab Output

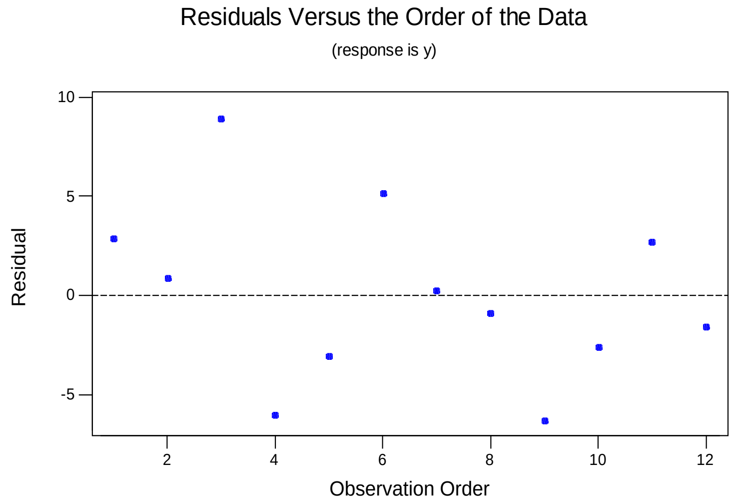
Regression Analysis: y versus x1, x2, x1^2, x2^2, x1x2

```
The regression equation is
y = 24.4 - 38.0 x1 + 0.7 x2 + 35.0 x1^2 + 11.1 x2^2 - 9.99 x1x2

Predictor        Coef    SE Coef        T       P      VIF
Constant        24.41      26.59     0.92   0.394
x1             -38.03      40.45    -0.94   0.383     89.6
x2               0.72      11.69     0.06   0.953     52.1
x1^2            34.98      21.56     1.62   0.156    103.9
x2^2           11.066      3.158     3.50   0.013    104.7
x1x2           -9.986      8.742    -1.14   0.297    105.1
```

```
S = 6.042              R-Sq = 99.4%         R-Sq(adj) = 98.9%
PRESS = 1327.71        R-Sq(pred) = 96.24%
r Analysis of Variance

Source             DF         SS         MS         F        P
Regression          5    35092.6     7018.5    192.23    0.000
Residual Error      6      219.1       36.5
  Lack of Fit       3       91.1       30.4      0.71    0.607
  Pure Error        3      128.0       42.7
Total              11    35311.7

7 rows with no replicates

Source        DF     Seq SS
x1             1    11552.0
x2             1    22950.3
x1^2           1       21.9
x2^2           1      520.8
x1x2           1       47.6
```







10.13.

(a) Consider the quadratic regression model from Problem 10.12. Compute *t* statistics for each model parameter and comment on the conclusions that follow from the quantities.

Minitab Output

Predictor	Coef	SE Coef	T	P	VIF
Constant	24.41	26.59	0.92	0.394	
x1	-38.03	40.45	-0.94	0.383	89.6
x2	0.72	11.69	0.06	0.953	52.1
x1^2	34.98	21.56	1.62	0.156	103.9
x2^2	11.066	3.158	3.50	0.013	104.7
x1x2	-9.986	8.742	-1.14	0.297	105.1

x_2^2 is the only model parameter that is statistically significant with a *t*-value of 3.50. A logical model might also include x_2 to preserve model hierarchy.

(b) Use the extra sum of squares method to evaluate the value of the quadratic terms, x_1^2, x_2^2 and x_1x_2 to the model.

The extra sum of squares due to β_2 is

$$SS_R\left(\beta_2 \middle| \beta_0, \beta_1\right) = SS_R\left(\beta_0, \beta_1, \beta_2\right) - SS_R\left(\beta_0, \beta_1\right) = SS_R\left(\beta_1, \beta_2 \middle| \beta_0\right) - SS_R\left(\beta_1 \middle| \beta_0\right)$$

$SS_R\left(\beta_1, \beta_2 \middle| \beta_0\right)$ sum of squares of regression for the model in Problem 10.12 = 35092.6

$$SS_R\left(\beta_1 \middle| \beta_0\right) = 34502.3$$

$$SS_R\left(\beta_2 \middle| \beta_0, \beta_1\right) = 35092.6 - 34502.3 = 590.3$$

$$F_0 = \frac{SS_R\left(\beta_2 \middle| \beta_0, \beta_1\right)/3}{MS_E} = \frac{590.3/3}{36.511} = 5.3892$$

Since $F_{0.05,3,6} = 4.76$, then the addition of the quadratic terms to the model is significant. The *p*-values indicate that it's probably the term x_2^2 that is responsible for this.

10.14. ***Relationship between analysis of variance and regression.*** Any analysis of variance model can be expressed in terms of the general linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where the **X** matrix consists of zeros and ones. Show that the single-factor model $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$, *i*=1,2,3, *j*=1,2,3,4 can be written in general linear model form. Then

(a) Write the normal equations $(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$ and compare them with the normal equations found for the model in Chapter 3.

The normal equations are $(\mathbf{X}'\mathbf{X})\hat{\boldsymbol{\beta}} = \mathbf{X}'\mathbf{y}$

$$\begin{bmatrix} 12 & 4 & 4 & 4 \\ 4 & 4 & 0 & 0 \\ 4 & 0 & 4 & 0 \\ 4 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\tau}_3 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{bmatrix}$$

which are in agreement with the results of Chapter 3.

(b) Find the rank of $\mathbf{X}'\mathbf{X}$. Can $(\mathbf{X}'\mathbf{X})^{-1}$ be obtained?

$\mathbf{X}'\mathbf{X}$ is a 4 x 4 matrix of rank 3, because the last three columns add to the first column. Thus $(\mathbf{X}'\mathbf{X})^{-1}$ does not exist.

(c) Suppose the first normal equation is deleted and the restriction $\sum_{i=1}^{3} n\hat{\tau}_i = 0$ is added. Can the resulting system of equations be solved? If so, find the solution. Find the regression sum of squares $\hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y}$, and compare it to the treatment sum of squares in the single-factor model.

Imposing $\sum_{i=1}^{3} n\hat{\tau}_i = 0$ yields the normal equations

$$\begin{bmatrix} 0 & 4 & 4 & 4 \\ 4 & 4 & 0 & 0 \\ 4 & 0 & 4 & 0 \\ 4 & 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\tau}_1 \\ \hat{\tau}_2 \\ \hat{\tau}_3 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{bmatrix}$$

The solution to this set of equations is

$$\hat{\mu} = \frac{y_{..}}{12} = \bar{y}_{..}$$
$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$$

This solution was found by solving the last three equations for $\hat{\tau}_i$, yielding $\hat{\tau}_i = \bar{y}_{i.} - \hat{\mu}$, and then substituting in the first equation to find $\hat{\mu} = \bar{y}_{..}$

The regression sum of squares is

$$SS_R(\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{y} = \bar{y}_{..}y_{..} + \sum_{i=1}^{a}\left(\bar{y}_{i.} - y_{..}\right)^2 = \frac{y_{..}^2}{an} + \sum_{i=1}^{a}\frac{\bar{y}_{i.}^2}{n} - \frac{y_{..}^2}{an} = \sum_{i=1}^{a}\frac{\bar{y}_{i.}^2}{n}$$

with a degrees of freedom. This is the same result found in Chapter 3. For more discussion of the relationship between analysis of variance and regression, see Montgomery and Peck (1992).

10.15. Suppose that we are fitting a straight line and we desire to make the variance of $\hat{\beta}_1$ as small as possible. Restricting ourselves to an even number of experimental points, where should we place these points so as to minimize $V\left(\hat{\beta}_1\right)$? (Note: Use the design called for in this exercise with great caution because, even though it minimizes $V\left(\hat{\beta}_1\right)$, it has some undesirable properties; for example, see Myers and Montgomery (1995). Only if you are *very sure* the true functional relationship is linear should you consider using this design.

Since $V\left(\hat{\beta}_1\right) = \dfrac{\sigma^2}{S_{xx}}$, we may minimize $V\left(\hat{\beta}_1\right)$by making S_{xx} as large as possible. S_{xx} is maximized by spreading out the x_j's as much as possible. The experimenter usually has a "region of interest" for x. If n is even, $n/2$ of the observations should be run at each end of the "region of interest". If n is odd, then run one of the observations in the center of the region and the remaining $(n-1)/2$ at either end.

10.16. ***Weighted least squares.*** Suppose that we are fitting the straight line $y = \beta_0 + \beta_1 x + \varepsilon$, but the variance of the y's now depends on the level of x; that is,

$$V\left(y|x_i\right) = \sigma_i^2 = \frac{\sigma^2}{w_i}, i = 1, 2, ..., n$$

where the w_i are known constants, often called weights. Show that if we choose estimates of the regression coefficients to minimize the weighted sum of squared errors given by $\sum_{i=1}^{n} w_i\left(y_i - \beta_0 + \beta_1 x_i\right)^2$, the resulting least squares normal equations are

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i + \hat{\beta}_1 \sum_{i=1}^{n} w_i x_i = \sum_{i=1}^{n} w_i y_i$$

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i x_i + \hat{\beta}_1 \sum_{i=1}^{n} w_i x_i^2 = \sum_{i=1}^{n} w_i x_i y_i$$

The least squares normal equations are found:

$$L = \sum_{i=1}^{n} \left(y_i - \beta_0 - \beta_1 x_1\right)^2 w_i$$
$$\frac{\partial L}{\partial \beta_0} = -2\sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_1\right) w_i = 0$$
$$\frac{\partial L}{\partial \beta_1} = -2\sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_1\right) x_1 w_i = 0$$

which simplify to

$$\hat{\beta}_0 \sum_{i=1}^{n} w_i + \hat{\beta}_1 \sum_{i=1}^{n} x_1 w_i = \sum_{i=1}^{n} w_i y_i$$
$$\hat{\beta}_0 \sum_{i=1}^{n} x_1 w_i + \hat{\beta}_1 \sum_{i=1}^{n} x_1^2 w_i = \sum_{i=1}^{n} w_i x_1 y_i$$

10.17. Consider the 2_{IV}^{4-1} design discussed in Example 10.5.

(a) Suppose you elect to augment the design with the single run selected in that example. Find the variances and covariances of the regression coefficients in the model (ignoring blocks):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{34} x_3 x_4 + \varepsilon$$

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 9 & -1 & -1 & -1 & 1 & 1 & -1 \\ -1 & 9 & 1 & 1 & -1 & -1 & 1 \\ -1 & 1 & 9 & 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & 9 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 9 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 9 & 7 \\ -1 & 1 & 1 & 1 & -1 & 7 & 9 \end{bmatrix}$$

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 0.1250 & 0 & 0 & 0 & 0 & -0.0625 & 0.0625 \\ 0 & 0.1250 & 0 & 0 & 0 & 0.0625 & -0.0625 \\ 0 & 0 & 0.1250 & 0 & 0 & 0.0625 & -0.0625 \\ 0 & 0 & 0 & 0.1250 & 0 & 0.0625 & -0.0625 \\ 0 & 0 & 0 & 0 & 0.1250 & -0.0625 & 0.0625 \\ -0.0625 & 0.0625 & 0.0625 & 0.0625 & -0.0625 & 0.4375 & -0.3750 \\ 0.0625 & -0.0625 & -0.0625 & -0.0625 & 0.0625 & -0.3750 & 0.4375 \end{bmatrix}$$

(b) Are there any other runs in the alternate fraction that would de-alias *AB* from *CD*?

Any other run from the alternate fraction will de-alias *AB* from *CD.*

(c) Suppose you augment the design with four runs suggested in Example 10.5. Find the variance and the covariances of the regression coefficients (ignoring blocks) for the model in part (a).

Choose 4 runs that are one of the quarter fractions not used in the principal half fraction.

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

$$\mathbf{X'X} = \begin{bmatrix} 12 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 12 & 4 & 0 & 0 & 0 \\ 0 & 0 & 4 & 12 & 0 & 0 & 0 \\ 0 & -4 & 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 12 & 4 \\ 0 & 0 & 0 & 0 & 0 & 4 & 12 \end{bmatrix}$$

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 0.08333 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.09375 & 0 & 0 & 0.03125 & 0 & 0 \\ 0 & 0 & 0.09375 & -0.03125 & 0 & 0 & 0 \\ 0 & 0 & -0.03125 & 0.09375 & 0 & 0 & 0 \\ 0 & 0.03125 & 0 & 0 & 0.09375 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.09375 & -0.03125 \\ 0 & 0 & 0 & 0 & 0 & -0.03125 & 0.09375 \end{bmatrix}$$

(d) Considering parts (a) and (c), which augmentation strategy would you prefer and why?

If you only have the resources to run one more run, then choose the one-run augmentation. But if resources are not scarce, then augment the design in multiples of two runs, to keep the design orthogonal. By using four runs, smaller variances of the regression coefficients are achieved along with a simpler covariance structure.

10.18. Consider the 2_{III}^{7-4}. Suppose after running the experiment, the largest observed effects are *A* + *BD*, *B* + *AD*, and *D* + *AB*. You wish to augment the original design with a group of four runs to de-alias these effects.

(a) Which four runs would you make?

Take the first four runs of the original experiment and change the sign on *A*.

Design Expert Output

			Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Std	Run	Block	A:x1	B:x2	C:x3	D:x4	E:x5	F:x6	G:x7
1	1	Block 1	-1	-1	-1	1	1	1	-1
2	2	Block 1	1	-1	-1	-1	-1	1	1
3	3	Block 1	-1	1	-1	-1	1	-1	1
4	4	Block 1	1	1	-1	1	-1	-1	-1
5	5	Block 1	-1	-1	1	1	-1	-1	1
6	6	Block 1	1	-1	1	-1	1	-1	-1
7	7	Block 1	-1	1	1	-1	-1	1	-1
8	8	Block 1	1	1	1	1	1	1	1
9	9	Block 2	1	-1	-1	1	1	1	-1
10	10	Block 2	-1	-1	-1	-1	-1	1	1
11	11	Block 2	1	1	-1	-1	1	-1	1
12	12	Block 2	-1	1	-1	1	-1	-1	-1

Main effects and interactions of interest are:

x1	x2	x4	x1x2	x1x4	x2x4
-1	-1	1	1	-1	-1
1	-1	-1	-1	-1	1
-1	1	-1	-1	1	-1
1	1	1	1	1	1
-1	-1	1	1	-1	-1
1	-1	-1	-1	-1	1
-1	1	-1	-1	1	-1
1	1	1	1	1	1
1	**-1**	**1**	**-1**	**1**	**-1**
-1	**-1**	**-1**	**1**	**1**	**1**
1	**1**	**-1**	**1**	**-1**	**-1**
-1	**1**	**1**	**-1**	**-1**	**1**

(b) Find the variances and covariances of the regression coefficients in the model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_4 x_4 + \beta_{12} x_1 x_2 + \beta_{14} x_1 x_4 + \beta_{24} x_2 x_4 + \varepsilon$$

$$\mathbf{X'X} = \begin{bmatrix} 12 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 12 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 12 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 12 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 12 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 12 \end{bmatrix}$$

$$(\mathbf{X'X})^{-1} = \begin{bmatrix} 0.08333 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.09375 & 0 & 0 & 0 & 0 & -0.03125 \\ 0 & 0 & 0.09375 & 0 & 0 & -0.03125 & 0 \\ 0 & 0 & 0 & 0.09375 & -0.03125 & 0 & 0 \\ 0 & 0 & 0 & -0.03125 & 0.09375 & 0 & 0 \\ 0 & 0 & -0.03125 & 0 & 0 & 0.09375 & 0 \\ 0 & -0.03125 & 0 & 0 & 0 & 0 & 0.09375 \end{bmatrix}$$

(c) Is it possible to de-alias these effects with fewer than four additional runs?

It is possible to de-alias these effects in only two runs. By utilizing *Design Expert's* design augmentation *D*-optimal factorial function, the runs 9 and 10 (Block 2) were generated as follows:

Design Expert Output

			Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Std	Run	Block	A:x1	B:x2	C:x3	D:x4	E:x5	F:x6	G:x7
1	1	Block 1	-1	-1	-1	1	1	1	-1
2	2	Block 1	1	-1	-1	-1	-1	1	1
3	3	Block 1	-1	1	-1	-1	1	-1	1
4	4	Block 1	1	1	-1	1	-1	-1	-1
5	5	Block 1	-1	-1	1	1	-1	-1	1
6	6	Block 1	1	-1	1	-1	1	-1	-1
7	7	Block 1	-1	1	1	-1	-1	1	-1
8	8	Block 1	1	1	1	1	1	1	1
9	9	Block 2	1	-1	-1	1	-1	-1	-1
10	10	Block 2	-1	-1	-1	-1	-1	-1	-1

Chapter 11
Response Surface Methods and Designs
Solutions

11.1. A chemical plant produces oxygen by liquefying air and separating it into its component gases by fractional distillation. The purity of the oxygen is a function of the main condenser temperature and the pressure ratio between the upper and lower columns. Current operating conditions are temperature $(\xi_1) =$ -220°C and pressure ratio $(\xi_2) = 1.2$. Using the following data find the path of steepest ascent.

Temperature (ξ_1)	Pressure Ratio (ξ_2)	Purity
-225	1.1	82.8
-225	1.3	83.5
-215	1.1	84.7
-215	1.3	85.0
-220	1.2	84.1
-220	1.2	84.5
-220	1.2	83.9
-220	1.2	84.3

Design Expert Output

Response: Purity
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	3.14	2	1.57	26.17	0.0050	significant
A	*2.89*	*1*	*2.89*	*48.17*	*0.0023*	
B	*0.25*	*1*	*0.25*	*4.17*	*0.1108*	
Curvature	0.080	1	0.080	1.33	0.3125	not significant
Residual	0.24	4	0.060			
Lack of Fit	*0.040*	*1*	*0.040*	*0.60*	*0.4950*	*not significant*
Pure Error	*0.20*	*3*	*0.067*			
Cor Total	3.46	7				

The Model F-value of 26.17 implies the model is significant. There is only
a 0.50% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.24	R-Squared	0.9290
Mean	84.10	Adj R-Squared	0.8935
C.V.	0.29	Pred R-Squared	0.7123
PRESS	1.00	Adeq Precision	12.702

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	84.00	1	0.12	83.66	84.34	
A-Temperature	0.85	1	0.12	0.51	1.19	1.00
B-Pressure Ratio	0.25	1	0.12	-0.090	0.59	1.00
Center Point	0.20	1	0.17	-0.28	0.68	1.00

Final Equation in Terms of Coded Factors:

Purity =
+84.00
+0.85 * A
+0.25 * B

Final Equation in Terms of Actual Factors:

```
       Purity  =
   +118.40000
     +0.17000   * Temperature
     +2.50000   * Pressure Ratio
```

From the computer output use the model $\hat{y} = 84 + 0.85x_1 + 0.25x_2$ as the equation for steepest ascent. Suppose we use a one degree change in temperature as the basic step size. Thus, the path of steepest ascent passes through the point (x_1=0, x_2=0) and has a slope 0.25/0.85. In the coded variables, one degree of temperature is equivalent to a step of $\Delta x_1 = 1/5=0.2$. Thus, $\Delta x_2 = (0.25/0.85)0.2=0.059$. The path of steepest ascent is:

	Coded	Variables	Natural	Variables
	x_1	x_2	ξ_1	ξ_2
Origin	0	0	-220	1.2
Δ	0.2	0.059	1	0.0059
Origin + Δ	0.2	0.059	-219	1.2059
Origin +5 Δ	1.0	0.295	-215	1.2295
Origin +7 Δ	1.40	0.413	-213	1.2413

11.2. An industrial engineer has developed a computer simulation model of a two-item inventory system. The decision variables are the order quantity and the reorder point for each item. The response to be minimized is the total inventory cost. The simulation model is used to produce the data shown in the Table P11.1. Identify the experimental design. Find the path of steepest descent.

Table P11.1

Item 1		Item 2		
Order Quantity (x1)	Reorder Point (x2)	Order Quantity (x3)	Reorder Point (x4)	Total Cost
100	25	250	40	625
140	45	250	40	670
140	25	300	40	663
140	25	250	80	654
100	45	300	40	648
100	45	250	80	634
100	25	300	80	692
140	45	300	80	686
120	35	275	60	680
120	35	275	60	674
120	35	275	60	681

The design is a 2^{4-1} fractional factorial with generator *I=ABCD*, and three center points.

Design Expert Output

Response: Total Cost
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	3880.00	6	646.67	63.26	0.0030	significant
A	*684.50*	*1*	*684.50*	*66.96*	*0.0038*	
C	*1404.50*	*1*	*1404.50*	*137.40*	*0.0013*	
D	*450.00*	*1*	*450.00*	*44.02*	*0.0070*	
AC	*392.00*	*1*	*392.00*	*38.35*	*0.0085*	
AD	*264.50*	*1*	*264.50*	*25.88*	*0.0147*	
CD	*684.50*	*1*	*684.50*	*66.96*	*0.0038*	

Curvature	815.52	1	815.52	79.78	0.0030	significant
Residual	30.67	3	10.22			
Lack of Fit	*2.00*	*1*	*2.00*	*0.14*	*0.7446*	*not significant*
Pure Error	*28.67*	*2*	*14.33*			
Cor Total	4726.18	10				

The Model F-value of 63.26 implies the model is significant. There is only a 0.30% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	3.20	R-Squared	0.9922
Mean	664.27	Adj R-Squared	0.9765
C.V.	0.48	Pred R-Squared	0.9593
PRESS	192.50	Adeq Precision	24.573

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	659.00	1	1.13	655.40	662.60	
A-Item 1 QTY	9.25	1	1.13	5.65	12.85	1.00
C-Item 2 QTY	13.25	1	1.13	9.65	16.85	1.00
D-Item 2 Reorder	7.50	1	1.13	3.90	11.10	1.00
AC	-7.00	1	1.13	-10.60	-3.40	1.00
AD	-5.75	1	1.13	-9.35	-2.15	1.00
CD	9.25	1	1.13	5.65	12.85	1.00
Center Point	19.33	1	2.16	12.44	26.22	1.00

Final Equation in Terms of Coded Factors:

Total Cost	=
+659.00	
+9.25	* A
+13.25	* C
+7.50	* D
-7.00	* A * C
-5.75	* A * D
+9.25	* C * D

Final Equation in Terms of Actual Factors:

Total Cost	=
+175.00000	
+5.17500	* Item 1 QTY
+1.10000	* Item 2 QTY
-2.98750	* Item 2 Reorder
-0.014000	* Item 1 QTY * Item 2 QTY
-0.014375	* Item 1 QTY * Item 2 Reorder
+0.018500	* Item 2 QTY * Item 2 Reorder

The equation used to compute the path of steepest ascent is $\hat{y} = 659 + 9.25x_1 + 13.25x_3 + 7.50x_4$. Notice that even though the model contains interaction, it is relatively common practice to ignore the interactions in computing the path of steepest ascent. This means that the path constructed is only an approximation to the path that would have been obtained if the interactions were considered, but it's usually close enough to give satisfactory results.

It is helpful to give a general method for finding the path of steepest ascent. Suppose we have a first-order model in *k* variables, say

$$\hat{y} = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i$$

The path of steepest ascent passes through the origin, **x**=**0**, and through the point on a hypersphere of radius, *R* where $\hat{y}$ is a maximum. Thus, the *x*'s must satisfy the constraint

$$\sum_{i=1}^{k} x_i^2 = R^2$$

To find the set of x's that maximize $\hat{y}$ subject to this constraint, we maximize

$$L = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i - \lambda \left[\sum_{i=1}^{k} x_i^2 - R^2 \right]$$

where λ is a LaGrange multiplier. From $\partial L / \partial x_i = \partial L / \partial \lambda = 0$, we find

$$x_i = \frac{\hat{\beta}_i}{2\lambda}$$

It is customary to specify a basic step size in one of the variables, say Δx_j, and then calculate 2λ as $2\lambda = \hat{\beta}_j / \Delta x_j$. Then this value of 2λ can be used to generate the remaining coordinates of a point on the path of steepest ascent.

We demonstrate using the data from this problem. Suppose that we use -10 units in ξ_1 as the basic step size. Note that a decrease in ξ_1 is called for, because we are looking for a path of steepest decent. Now -10 units in ξ_1 is equal to -10/20 = -0.5 units change in x_1.

Thus, $2\lambda = \hat{\beta}_1 / \Delta x_1 = 9.25/(-0.5) = -18.50$

Consequently,

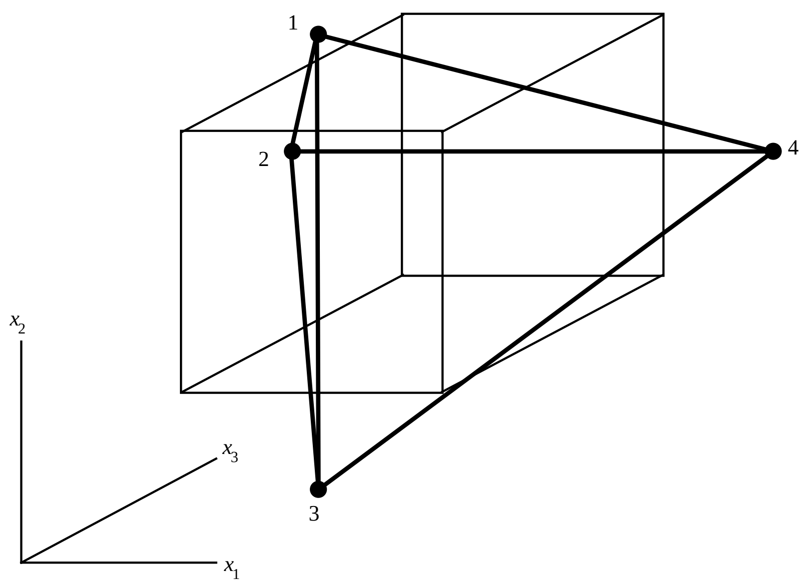
$$\Delta x_3 = \frac{\hat{\beta}_3}{2\lambda} = \frac{13.25}{-18.50} = -0.716$$
$$\Delta x_4 = \frac{\hat{\beta}_4}{2\lambda} = \frac{7.50}{-18.50} = -0.405$$

are the remaining coordinates of points along the path of steepest decent, in terms of the coded variables. The path of steepest decent is shown below:

	Coded Variables				Natural Variables			
	x_1	x_2	x_3	x_4	ξ_1	ξ_2	ξ_3	ξ_4
Origin	0	0	0	0	120	35	275	60
Δ	-0.50	0	-0.716	-0.405	-10	0	-17.91	-8.11
Origin + Δ	-0.50	0	-0.716	-0.405	110	35	257.09	51.89
Origin +2 Δ	-1.00	0	-1.432	-0.810	100	35	239.18	43.78

11.3. Verify that the following design is a simplex. Fit the first-order model and find the path of steepest ascent.

Position	x_1	x_2	x_3	y
1	0	$\sqrt{2}$	-1	18.5
2	$-\sqrt{2}$	0	1	19.8
3	0	$-\sqrt{2}$	-1	17.4
4	$\sqrt{2}$	0	1	22.5



The graphical representation of the design identifies a tetrahedron; therefore, the design is a simplex.

Design Expert Output

Response: **y**
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	14.49	3	4.83		
A	*3.64*	*1*	*3.64*		
B	*0.61*	*1*	*0.61*		
C	*10.24*	*1*	*10.24*		
Pure Error	0.000	0			
Cor Total	14.49	3			

Std. Dev.		R-Squared	1.0000
Mean	19.55	Adj R-Squared	
C.V.		Pred R-Squared	N/A
PRESS	N/A	Adeq Precision	0.000

Case(s) with leverage of 1.0000: Pred R-Squared and PRESS statistic not defined

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	19.55	1				
A-x1	1.35	1				1.00
B-x2	0.55	1				1.00
C-x3	1.60	1				1.00

Final Equation in Terms of Coded Factors:

y	=
+19.55	
+1.35	* A
+0.55	* B
+1.60	* C

Final Equation in Terms of Actual Factors:

y	=
+19.55000	
+0.95459	* x1
+0.38891	* x2
+1.60000	* x3

The first order model is $\hat{y} = 19.55 + 1.35x_1 + 0.55x_2 + 1.60x_3$.

To find the path of steepest ascent, let the basic step size be $\Delta x_3 = 1$. Then using the results obtained in the previous problem, we obtain

$$\Delta x_3 = \frac{\hat{\beta}_3}{2\lambda} \quad \text{or} \quad 1.0 = \frac{1.60}{2\lambda}$$

which yields $2\lambda = 1.60$. Then the coordinates of points on the path of steepest ascent are defined by

$$\Delta x_1 = \frac{\hat{\beta}_1}{2\lambda} = \frac{1.35}{1.60} = 0.84$$
$$\Delta x_2 = \frac{\hat{\beta}_2}{2\lambda} = \frac{0.55}{1.60} = 0.34$$

Therefore, in the coded variables we have:

	Coded x_1	Variables x_2	x_3
Origin	0	0	0
Δ	0.84	0.34	1.00
Origin + Δ	0.84	0.34	1.00
Origin +2 Δ	1.68	0.68	2.00

11.4. For the first-order model $\hat{y} = 60 + 1.5x_1 - 0.8x_2 + 2.0x_3$ find the path of steepest ascent. The variables are coded as $-1 \le x_i \le 1$.

Let the basic step size be $\Delta x_3 = 1$. $\Delta x_3 = \dfrac{\hat{\beta}_3}{2\lambda}$ or $1.0 = \dfrac{2.0}{2\lambda}$. Then $2\lambda = 2.0$

$$\Delta x_1 = \frac{\hat{\beta}_1}{2\lambda} = \frac{1.50}{2.0} = 0.75$$
$$\Delta x_2 = \frac{\hat{\beta}_2}{2\lambda} = \frac{-0.8}{2.0} = -0.40$$

Therefore, in the coded variables we have

	Coded	Variables	
	x_1	x_2	x_3
Origin	0	0	0
Δ	0.75	-0.40	1.00
Origin + Δ	0.75	-0.40	1.00
Origin +2 Δ	1.50	-0.80	2.00

11.5. The region of experimentation for three factors are time ($40 \le T_1 \le 80$ min), temperature ($200 \le T_2 \le 300$ °C), and pressure ($20 \le P \le 50$ psig). A first-order model in coded variables has been fit to yield data from a 2^3 design. The model is

$$\hat{y} = 30 + 5x_1 + 2.5x_2 + 3.5x_3$$

Is the point T_1 = 85, T_2 = 325, P=60 on the path of steepest ascent?

The coded variables are found with the following:

$$x_1 = \frac{T_1 - 60}{20} \qquad x_2 = \frac{T_2 - 250}{50} \qquad x_3 \frac{P_1 - 35}{15}$$

$$\Delta T_1 = 5 \quad \Delta x_1 = \frac{5}{20} = 0.25$$

$$\Delta x_1 = \frac{\hat{\beta}_1}{2\lambda} \text{ or } 0.25 = \frac{5}{2\lambda} \text{ where } 2\lambda = 20$$

$$\Delta x_2 = \frac{\hat{\beta}_2}{2\lambda} = \frac{2.5}{20} = 0.125$$

$$\Delta x_3 = \frac{\hat{\beta}_3}{2\lambda} = \frac{3.5}{20} = 0.175$$

	Coded	Variables		Natural	Variables	
	x_1	x_2	x_3	T_1	T_2	P
Origin	0	0	0	60	250	35
Δ	0.25	0.125	0.175	5	6.25	2.625
Origin + Δ	0.25	0.125	0.175	65	256.25	37.625
Origin +5 Δ	1.25	0.625	0.875	85	281.25	48.125

The point T_1=85, T_2=325, and P=60 is not on the path of steepest ascent.

11.6. The region of experimentation for two factors are temperature ($100 \le T \le 300°$ F) and catalyst feed rate ($10 \le C \le 30$ lb/h). A first order model in the usual ± 1 coded variables has been fit to a molecular weight response, yielding the following model.

$$\hat{y} = 2000 + 125x_1 + 40x_2$$

(a) Find the path of steepest ascent.

$$x_1 = \frac{T - 200}{100} \qquad x_2 = \frac{C - 20}{10}$$

$$\Delta T = 100 \qquad \Delta x_1 = \frac{100}{100} = 1$$

$$\Delta x_1 = \frac{\hat{\beta}_1}{2\lambda} \text{ or } 1 = \frac{125}{2\lambda} \qquad 2\lambda = 125$$

$$\Delta x_2 = \frac{\hat{\beta}_2}{2\lambda} = \frac{40}{125} = 0.32$$

	Coded x_1	Variables x_2	Natural T	Variables C
Origin	0	0	200	20
Δ	1	0.32	100	3.2
Origin + Δ	1	0.32	300	23.2
Origin +5 Δ	5	1.60	700	36.0

(b) It is desired to move to a region where molecular weights are above 2500. Based on the information you have from the experimentation in this region, about how may steps along the path of steepest ascent might be required to move to the region of interest?

$$\Delta\hat{y} = \Delta x_1 \hat{\beta}_1 + \Delta x_2 \hat{\beta}_2 = (1)(125) + (0.32)(40) = 137.8$$

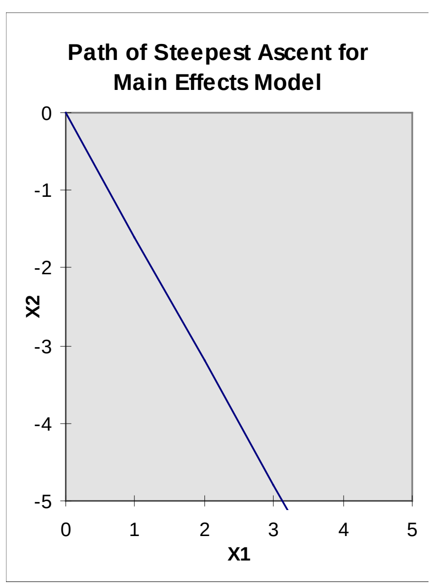
$$\#\,Steps = \frac{2500 - 2000}{137.8} = 3.63 \rightarrow 4$$

11.7. The path of steepest ascent is usually computed assuming that the model is truly first-order.; that is, there is no interaction. However, even if there is interaction, steepest ascent ignoring the interaction still usually produces good results. To illustrate, suppose that we have fit the model

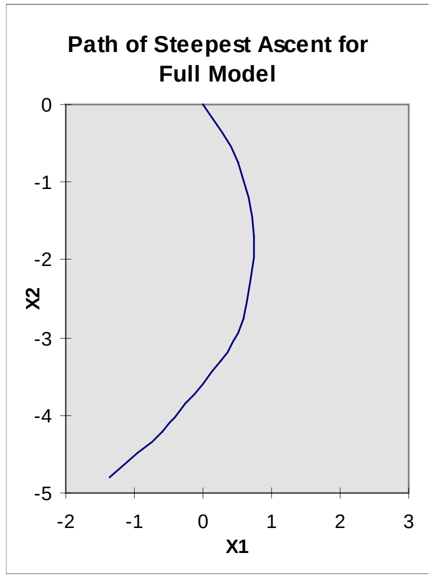
$$\hat{y} = 20 + 5x_1 - 8x_2 + 3x_1x_2$$

using coded variables (-1 ≤ x_1 ≤ +1)

(a) Draw the path of steepest ascent that you would obtain if the interaction were ignored.



(b) Draw the path of steepest ascent that you would obtain with the interaction included in the model. Compare this with the path found in part (a).



11.8. The data shown in Table P11.2 were collected in an experiment to optimize crystal growth as a function of three variables x_1, x_2, and x_3. Large values of y (yield in grams) are desirable. Fit a second order model and analyze the fitted surface. Under what set of conditions is maximum growth achieved?

Table P11.2

x_1	x_2	x_3	y
-1	-1	-1	66
-1	-1	1	70
-1	1	-1	78
-1	1	1	60
1	-1	-1	80
1	-1	1	70
1	1	-1	100
1	1	1	75
-1.682	0	0	100
1.682	0	0	80
0	-1.682	0	68
0	1.682	0	63
0	0	-1.682	65
0	0	1.682	82
0	0	0	113
0	0	0	100
0	0	0	118
0	0	0	88
0	0	0	100
0	0	0	85

Design Expert Output

Response: Yield

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	3662.00	9	406.89	2.19	0.1194	not significant

A	*22.08*	*1*	*22.08*	*0.12*	*0.7377*	
B	*25.31*	*1*	*25.31*	*0.14*	*0.7200*	
C	*30.50*	*1*	*30.50*	*0.16*	*0.6941*	
A^2	*204.55*	*1*	*204.55*	*1.10*	*0.3191*	
B^2	*2226.45*	*1*	*2226.45*	*11.96*	*0.0061*	
C^2	*1328.46*	*1*	*1328.46*	*7.14*	*0.0234*	
AB	*66.12*	*1*	*66.12*	*0.36*	*0.5644*	
AC	*55.13*	*1*	*55.13*	*0.30*	*0.5982*	
BC	*171.13*	*1*	*171.13*	*0.92*	*0.3602*	
Residual	1860.95	10	186.09			
Lack of Fit	*1001.61*	*5*	*200.32*	*1.17*	*0.4353*	*not significant*
Pure Error	*859.33*	*5*	*171.87*			
Cor Total	5522.95	19				

The "Model F-value" of 2.19 implies the model is not significant relative to the noise. There is a 11.94 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	13.64	R-Squared	0.6631
Mean	83.05	Adj R-Squared	0.3598
C.V.	16.43	Pred R-Squared	-0.6034
PRESS	8855.23	Adeq Precision	3.882

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	100.67	1	5.56	88.27	113.06	
A-x1	1.27	1	3.69	-6.95	9.50	1.00
B-x2	1.36	1	3.69	-6.86	9.59	1.00
C-x3	-1.49	1	3.69	-9.72	6.73	1.00
A^2	-3.77	1	3.59	-11.77	4.24	1.02
B^2	-12.43	1	3.59	-20.44	-4.42	1.02
C^2	-9.60	1	3.59	-17.61	-1.59	1.02
AB	2.87	1	4.82	-7.87	13.62	1.00
AC	-2.63	1	4.82	-13.37	8.12	1.00
BC	-4.63	1	4.82	-15.37	6.12	1.00

Final Equation in Terms of Coded Factors:

Yield	=
+100.67	
+1.27	* A
+1.36	* B
-1.49	* C
-3.77	* A^2
-12.43	* B^2
-9.60	* C^2
+2.87	* A * B
-2.63	* A * C
-4.63	* B * C

Final Equation in Terms of Actual Factors:

Yield	=
+100.66609	
+1.27146	* x1
+1.36130	* x2
-1.49445	* x3
-3.76749	* $x1^2$
-12.42955	* $x2^2$
-9.60113	* $x3^2$
+2.87500	* x1 * x2
-2.62500	* x1 * x3
-4.62500	* x2 * x3

There are so many non-significant terms in this model that we should consider eliminating some of them. A reasonable reduced model is shown below.

Design Expert Output

Response: Yield

ANOVA for Response Surface Reduced Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	3143.00	4	785.75	4.95	0.0095	significant
B	25.31	1	25.31	0.16	0.6952	
C	30.50	1	30.50	0.19	0.6673	
B2	2115.31	1	2115.31	13.33	0.0024	
C2	1239.17	1	1239.17	7.81	0.0136	
Residual	2379.95	15	158.66			
Lack of Fit	*1520.62*	*10*	*152.06*	*0.88*	*0.5953*	*not significant*
Pure Error	*859.33*	*5*	*171.87*			
Cor Total	5522.95	19				

The Model F-value of 4.95 implies the model is significant. There is only
a 0.95% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	12.60	R-Squared	0.5691
Mean	83.05	Adj R-Squared	0.4542
C.V.	15.17	Pred R-Squared	0.1426
PRESS	4735.52	Adeq Precision	5.778

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	97.58	1	4.36	88.29	106.88	
B-x2	1.36	1	3.41	-5.90	8.63	1.00
C-x3	-1.49	1	3.41	-8.76	5.77	1.00
B2	-12.06	1	3.30	-19.09	-5.02	1.01
C2	-9.23	1	3.30	-16.26	-2.19	1.01

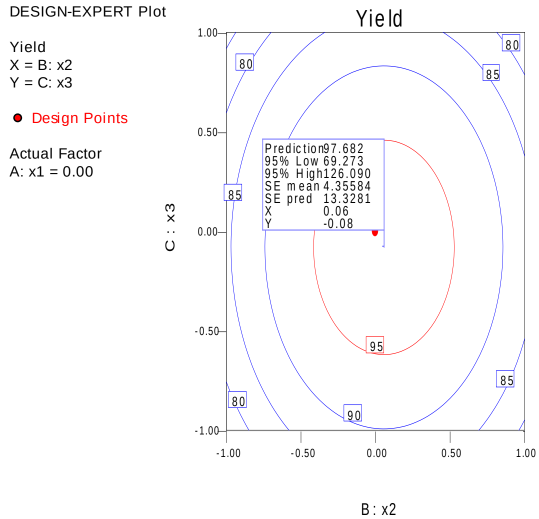
Final Equation in Terms of Coded Factors:

$$\begin{aligned} \text{Yield} &= \\ &+97.58 \\ &+1.36 \quad * B \\ &-1.49 \quad * C \\ &-12.06 \quad * B^2 \\ &-9.23 \quad * C^2 \end{aligned}$$

Final Equation in Terms of Actual Factors:

$$\begin{aligned} \text{Yield} &= \\ &+97.58260 \\ &+1.36130 \quad * x2 \\ &-1.49445 \quad * x3 \\ &-12.05546 \quad * x2^2 \\ &-9.22703 \quad * x3^2 \end{aligned}$$

The contour plot identifies a maximum near the center of the design space.



11.9. The data in Table P11.3 were collected by a chemical engineer. The response y is filtration time, x_1 is temperature, and x_2 is pressure. Fit a second-order model.

Table P11.3

x_1	x_2	y
-1	-1	54
-1	1	45
1	-1	32
1	1	47
-1.414	0	50
1.414	0	53
0	-1.414	47
0	1.414	51
0	0	41
0	0	39
0	0	44
0	0	42
0	0	40

Design Expert Output

Response: **y**

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	273.41	4	68.35	2.78	0.1018	not significant
A	*31.04*	*1*	*31.04*	*1.26*	*0.2937*	
B	*16.99*	*1*	*16.99*	*0.69*	*0.4299*	
A2	*81.39*	*1*	*81.39*	*3.31*	*0.1063*	
AB	*144.00*	*1*	*144.00*	*5.86*	*0.0418*	
Residual	196.59	8	24.57			
Lack of Fit	*181.79*	*4*	*45.45*	*12.28*	*0.0162*	*significant*
Pure Error	*14.80*	*4*	*3.70*			
Cor Total	470.00	12				

The "Model F-value" of 2.78 implies the model is not significant relative to the noise. There is a 10.18 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	4.96	R-Squared	0.5817
Mean	45.00	Adj R-Squared	0.3726

C.V.	11.02	Pred R-Squared	-0.7171
PRESS	807.05	Adeq Precision	5.185

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	42.91	1	1.79	38.78	47.04	
A-Temperature	-2.79	1	2.48	-8.50	2.93	1.00
B-Pressure	2.06	1	2.48	-3.65	7.78	1.00
A2	6.78	1	3.73	-1.81	15.38	1.00
AB	12.00	1	4.96	0.57	23.43	1.00

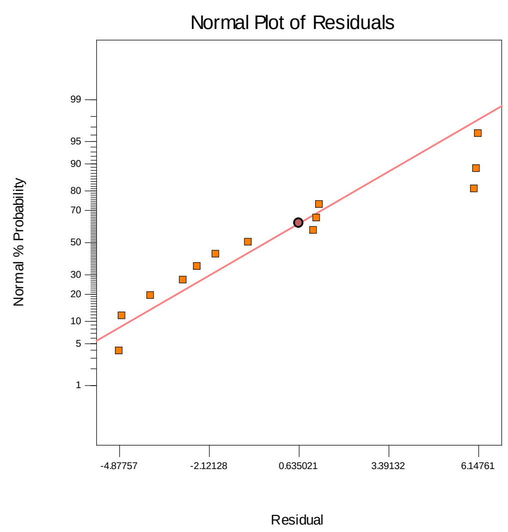
Final Equation in Terms of Coded Factors:

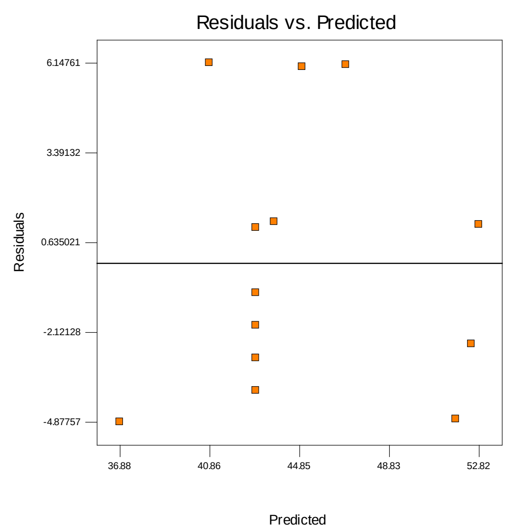
Time =
+42.91
+1.28 * A
-1.79 * B
+3.39 * A^2
+6.00 * A * B
Time =
+42.91
-2.79 * A
+2.06 * B
+6.78 * A2
+12.00 * A * B

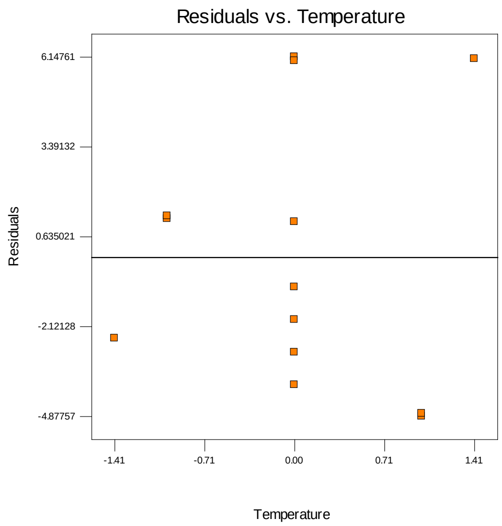
Final Equation in Terms of Actual Factors:

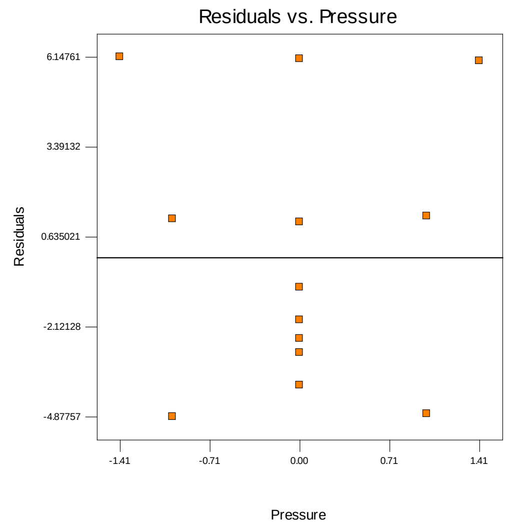
Time =
+42.91304
-1.96968 * Temperature
+1.45711 * Pressure
+3.39131 * Temperature2
+6.00000 * Temperature * Pressure

The lack of fit test in the above analysis is significant. The residual plots below are not unusual; however, the three largest residuals are axial points. Including cubic terms did not improve the lack of fit estimate and only made the PRESS larger.



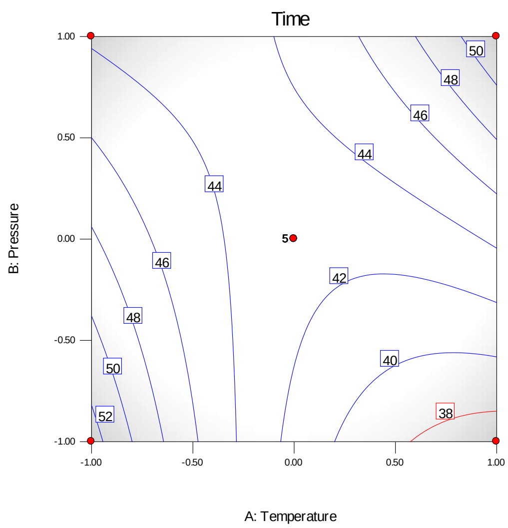


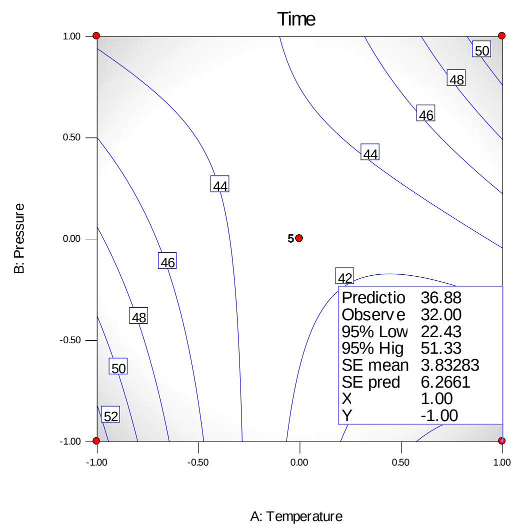




(a) What operating conditions would you recommend if the objective is to minimize the filtration time?

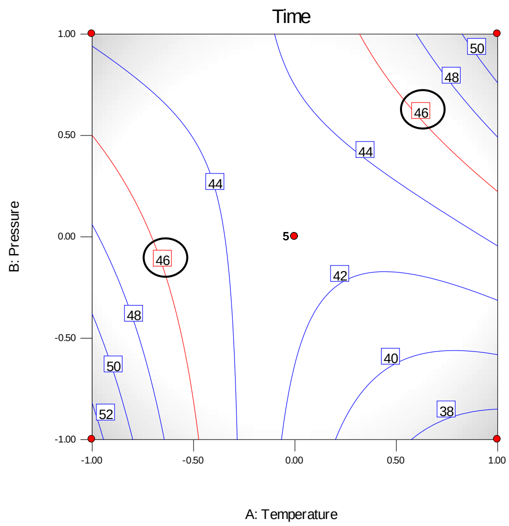
The contour plot below identifies the minimum filtration rate with the Temperature set at 1 and the Pressure at -1. By setting the flag at this location in *Design Expert*, the prediction of 36.88 filtration rate can be quickly estimated.





(b) What operating conditions would you recommend if the objective is to operate the process at a mean filtration time very close to 46?

The contour plot below identifies filtration time of 46 with the highlighted contours. There are two regions where a filtration rate of 46 is achievable. If higher temperature and pressure require higher operating costs, it would be preferential to chose the region with lower temperature and pressure.



11.10. The hexagon design in Table P11.4 is used in an experiment that has the objective of fitting a second-order model.

Table P11.4

x_1	x_2	y
1	0	68
0.5	$\sqrt{0.75}$	74
-0.5	$\sqrt{0.75}$	65
-1	0	60
-0.5	$-\sqrt{0.75}$	63
0.5	$-\sqrt{0.75}$	70
0	0	58
0	0	60
0	0	57
0	0	55
0	0	69

(a) Fit the second-order model.

Design Expert Output

Response: **y**

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	245.26	5	49.05	1.89	0.2500	not significant
A	*85.33*	*1*	*85.33*	*3.30*	*0.1292*	
B	*9.00*	*1*	*9.00*	*0.35*	*0.5811*	
A^2	*25.20*	*1*	*25.20*	*0.97*	*0.3692*	
B^2	*129.83*	*1*	*129.83*	*5.01*	*0.0753*	
AB	*1.00*	*1*	*1.00*	*0.039*	*0.8519*	
Residual	129.47	5	25.89			
Lack of Fit	*10.67*	*1*	*10.67*	*0.36*	*0.5813*	*not significant*
Pure Error	*118.80*	*4*	*29.70*			
Cor Total	374.73	10				

The "Model F-value" of 1.89 implies the model is not significant relative to the noise. There is a

25.00 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	5.09	R-Squared	0.6545
Mean	63.55	Adj R-Squared	0.3090
C.V.	8.01	Pred R-Squared	-0.5201
PRESS	569.63	Adeq Precision	3.725

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	59.80	1	2.28	53.95	65.65	
A-x1	5.33	1	2.94	-2.22	12.89	1.00
B-x2	1.73	1	2.94	-5.82	9.28	1.00
A^2	4.20	1	4.26	-6.74	15.14	1.00
B^2	9.53	1	4.26	-1.41	20.48	1.00
AB	1.15	1	5.88	-13.95	16.26	1.00

Final Equation in Terms of Coded Factors:

$$\begin{aligned} y &= \\ &+59.80 \\ &+5.33 \ * A \\ &+1.73 \ * B \\ &+4.20 \ * A^2 \\ &+9.53 \ * B^2 \\ &+1.15 \ * A * B \end{aligned}$$

(b) Perform the canonical analysis. What type of surface has been found?

The full quadratic model is used in the following analysis because the reduced model is singular.

Solution

Variable	**Critical Value**
X1	-0.627658
X2	-0.052829
Predicted Value at Solution	58.080492

Eigenvalues and Eigenvectors

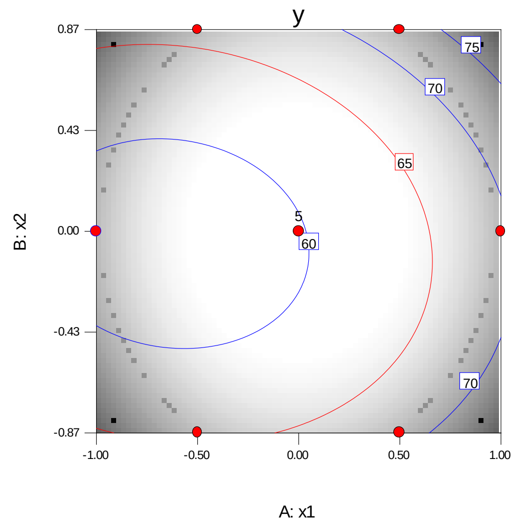
Variable	9.5957	4.1382
X1	0.10640	0.99432
X2	0.99432	-0.10640

Since both eigenvalues are positive, the response is a minimum at the stationary point.

(c) What operating conditions on x_1 and x_2 lead to the stationary point?

The stationary point is (x_1,x_2) = (-0.62766, -0.05283)

(d) Where would you run this process if the objective is to obtain a response that is as close to 65 as possible?



Any value of x_1 and x_2 that give a point on the contour with value of 65 would be satisfactory.

11.11. An experimenter has run a Box-Behnken design and has obtained the results as shown in Table P11.5, where the response variable is the viscosity of a polymer.

Table P11.5

Level	Temp.	Agitation Rate	Pressure	x_1	x_2	x_3
High	200	10.0	25	+1	+1	+1
Middle	175	7.5	20	0	0	0
Low	150	5.0	15	-1	-1	-1

Run	x_1	x_2	x_3	y_1
1	-1	-1	0	535
2	1	-1	0	580
3	-1	1	0	596
4	1	1	0	563
5	-1	0	-1	645
6	1	0	-1	458
7	-1	0	1	350
8	1	0	1	600
9	0	-1	-1	595
10	0	1	-1	648
11	0	-1	1	532
12	0	1	1	656
13	0	0	0	653
14	0	0	0	599
15	0	0	0	620

(a) Fit the second-order model.

Design Expert Output

Response: Viscosity

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	89652.58	9	9961.40	9.54	0.0115	significant
A	*703.12*	*1*	*703.12*	*0.67*	*0.4491*	
B	*6105.12*	*1*	*6105.12*	*5.85*	*0.0602*	
C	*5408.00*	*1*	*5408.00*	*5.18*	*0.0719*	
A^2	*20769.23*	*1*	*20769.23*	*19.90*	*0.0066*	
B^2	*1404.00*	*1*	*1404.00*	*1.35*	*0.2985*	
C^2	*4719.00*	*1*	*4719.00*	*4.52*	*0.0868*	
AB	*1521.00*	*1*	*1521.00*	*1.46*	*0.2814*	
AC	*47742.25*	*1*	*47742.25*	*45.74*	*0.0011*	
BC	*1260.25*	*1*	*1260.25*	*1.21*	*0.3219*	
Residual	5218.75	5	1043.75			
Lack of Fit	*3736.75*	*3*	*1245.58*	*1.68*	*0.3941*	*not significant*
Pure Error	*1482.00*	*2*	*741.00*			
Cor Total	94871.33	14				

The Model F-value of 9.54 implies the model is significant. There is only a 1.15% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	32.31	R-Squared	0.9450
Mean	575.33	Adj R-Squared	0.8460
C.V.	5.62	Pred R-Squared	0.3347
PRESS	63122.50	Adeq Precision	10.425

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	624.00	1	18.65	576.05	671.95	
A-Temp	9.37	1	11.42	-19.99	38.74	1.00
B-Agitation Rate	27.62	1	11.42	-1.74	56.99	1.00
C-Pressure	-26.00	1	11.42	-55.36	3.36	1.00
A^2	-75.00	1	16.81	-118.22	-31.78	1.01
B^2	19.50	1	16.81	-23.72	62.72	1.01
C^2	-35.75	1	16.81	-78.97	7.47	1.01
AB	-19.50	1	16.15	-61.02	22.02	1.00
AC	109.25	1	16.15	67.73	150.77	1.00
BC	17.75	1	16.15	-23.77	59.27	1.00

Final Equation in Terms of Coded Factors:

Viscosity =
+624.00
+9.37 * A
+27.62 * B
-26.00 * C
-75.00 * A^2
+19.50 * B^2
-35.75 * C^2
-19.50 * A * B
+109.25 * A * C
+17.75 * B * C

Final Equation in Terms of Actual Factors:

Viscosity =
-629.50000
+27.23500 * Temp
-9.55000 * Agitation Rate
-111.60000 * Pressure
-0.12000 * Temp^2
+3.12000 * Agitation Rate^2
-1.43000 * Pressure^2

-0.31200	* Temp * Agitation Rate
+0.87400	* Temp * Pressure
+1.42000	* Agitation Rate * Pressure

(b) Perform the canonical analysis. What type of surface has been found?

Solution

Variable	**Critical Value**
X1	2.1849596
X2	-0.871371
X3	2.7586015
Predicted Value at Solution	586.34437

Eigevalues and Eigevectors

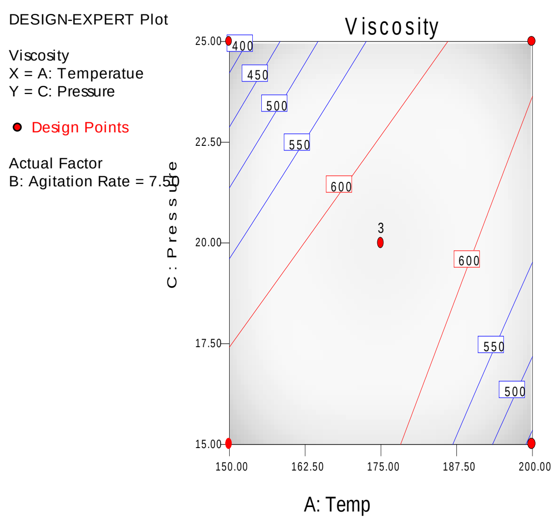
Variable	20.9229	2.5208	-114.694
X1	-0.02739	0.58118	0.81331
X2	0.99129	-0.08907	0.09703
X3	0.12883	0.80888	-0.57368

The system is a saddle point.

(c) What operating conditions on x_1, x_2, and x_3 lead to the stationary point?

The stationary point is (x_1, x_2, x_3) = (2.18496, -0.87167, 2.75860). This is outside the design region. It would be necessary to either examine contour plots or use numerical optimization methods to find desired operating conditions.

(d) What operating conditions would you recommend if it is important to obtain a viscosity that is as close to 600 as possible?



Any point on either of the contours showing a viscosity of 600 is satisfactory.

11.12. Consider the three-variable central composite design shown in Table P11.6. Analyze the data and draw conclusions, assuming that we wish to maximize conversion (y_1) with activity (y_2) between 55 and 60.

Table P11.6

Run	Time (min)	Temperature (°C)	Catalyst (%)	Conversion (%) y_1	Activity y_2
1	-1.000	-1.000	-1.000	74.00	53.20
2	1.000	-1.000	-1.000	51.00	62.90
3	-1.000	1.000	-1.000	88.00	53.40
4	1.000	1.000	-1.000	70.00	62.60
5	-1.000	-1.000	1.000	71.00	57.30
6	1.000	-1.000	1.000	90.00	67.90
7	-1.000	1.000	1.000	66.00	59.80
8	1.000	1.000	1.000	97.00	67.80
9	0.000	0.000	0.000	81.00	59.20
10	0.000	0.000	0.000	75.00	60.40
11	0.000	0.000	0.000	76.00	59.10
12	0.000	0.000	0.000	83.00	60.60
13	-1.682	0.000	0.000	76.00	59.10
14	1.682	0.000	0.000	79.00	65.90
15	0.000	-1.682	0.000	85.00	60.00
16	0.000	1.682	0.000	97.00	60.70
17	0.000	0.000	-1.682	55.00	57.40
18	0.000	0.000	1.682	81.00	63.20
19	0.000	0.000	0.000	80.00	60.80
20	0.000	0.000	0.000	91.00	58.90

Quadratic models are developed for the Conversion and Activity response variables as follows:

Design Expert Output

Response: Conversion
ANOVA for Response Surface Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	2555.73	9	283.97	12.76	0.0002	significant
A	*14.44*	*1*	*14.44*	*0.65*	*0.4391*	
B	*222.96*	*1*	*222.96*	*10.02*	*0.0101*	
C	*525.64*	*1*	*525.64*	*23.63*	*0.0007*	
A^2	*48.47*	*1*	*48.47*	*2.18*	*0.1707*	
B^2	*124.48*	*1*	*124.48*	*5.60*	*0.0396*	
C^2	*388.59*	*1*	*388.59*	*17.47*	*0.0019*	
AB	*36.13*	*1*	*36.13*	*1.62*	*0.2314*	
AC	*1035.13*	*1*	*1035.13*	*46.53*	*< 0.0001*	
BC	*120.12*	*1*	*120.12*	*5.40*	*0.0425*	
Residual	222.47	10	22.25			
Lack of Fit	*56.47*	*5*	*11.29*	*0.34*	*0.8692*	*not significant*
Pure Error	*166.00*	*5*	*33.20*			
Cor Total	287.28	19				

The Model F-value of 12.76 implies the model is significant. There is only a 0.02% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	4.72	R-Squared	0.9199
Mean	78.30	Adj R-Squared	0.8479
C.V.	6.02	Pred R-Squared	0.7566
PRESS	676.22	Adeq Precision	14.239

Coefficient **Standard** **95% CI** **95% CI**

Factor	Estimate	DF	Error	Low	High	VIF
Intercept	81.09	1	1.92	76.81	85.38	
A-Time	1.03	1	1.28	-1.82	3.87	1.00
B-Temperature	4.04	1	1.28	1.20	6.88	1.00
C-Catalyst	6.20	1	1.28	3.36	9.05	1.00
A2	-1.83	1	1.24	-4.60	0.93	1.02
B2	2.94	1	1.24	0.17	5.71	1.02
C2	-5.19	1	1.24	-7.96	-2.42	1.02
AB	2.13	1	1.67	-1.59	5.84	1.00
AC	11.38	1	1.67	7.66	15.09	1.00
BC	-3.87	1	1.67	-7.59	-0.16	1.00

Final Equation in Terms of Coded Factors:

Conversion	=
+81.09	
+1.03	* A
+4.04	* B
+6.20	* C
-1.83	* A2
+2.94	* B2
-5.19	* C2
+2.13	* A * B
+11.38	* A * C
-3.87	* B * C

Final Equation in Terms of Actual Factors:

Conversion	=
+81.09128	
+1.02845	* Time
+4.04057	* Temperature
+6.20396	* Catalyst
-1.83398	* Time2
+2.93899	* Temperature2
-5.19274	* Catalyst2
+2.12500	* Time * Temperature
+11.37500	* Time * Catalyst
-3.87500	* Temperature * Catalyst

Design Expert Output

Response: Activity

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	256.20	9	28.47	9.16	0.0009	significant
A	*175.35*	*1*	*175.35*	*56.42*	*< 0.0001*	
B	*0.89*	*1*	*0.89*	*0.28*	*0.6052*	
C	*67.91*	*1*	*67.91*	*21.85*	*0.0009*	
A^2	*10.05*	*1*	*10.05*	*3.23*	*0.1024*	
B^2	*0.081*	*1*	*0.081*	*0.026*	*0.8753*	
C^2	*0.047*	*1*	*0.047*	*0.015*	*0.9046*	
AB	*1.20*	*1*	*1.20*	*0.39*	*0.5480*	
AC	*0.011*	*1*	*0.011*	*3.620E-003*	*0.9532*	
BC	*0.78*	*1*	*0.78*	*0.25*	*0.6270*	
Residual	31.08	10	3.11			
Lack of Fit	*27.43*	*5*	*5.49*	*7.51*	*0.0226*	*significant*
Pure Error	*3.65*	*5*	*0.73*			
Cor Total	287.28	19				

The Model F-value of 9.16 implies the model is significant. There is only a 0.09% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.76	R-Squared	0.8918
Mean	60.51	Adj R-Squared	0.7945
C.V.	2.91	Pred R-Squared	0.2536
PRESS	214.43	Adeq Precision	10.911

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	59.85	1	0.72	58.25	61.45	
A-Time	3.58	1	0.48	2.52	4.65	1.00
B-Temperature	0.25	1	0.48	-0.81	1.32	1.00
C-Catalyst	2.23	1	0.48	1.17	3.29	1.00
A^2	0.83	1	0.46	-0.20	1.87	1.02
B^2	0.075	1	0.46	-0.96	1.11	1.02
C^2	0.057	1	0.46	-0.98	1.09	1.02
AB	-0.39	1	0.62	-1.78	1.00	1.00
AC	-0.038	1	0.62	-1.43	1.35	1.00
BC	0.31	1	0.62	-1.08	1.70	1.00

Final Equation in Terms of Coded Factors:

Conversion =
+59.85
+3.58 * A
+0.25 * B
+2.23 * C
+0.83 * A^2
+0.075 * B^2
+0.057 * C^2
-0.39 * A * B
-0.038 * A * C
+0.31 * B * C

Final Equation in Terms of Actual Factors:

Conversion =
+59.84984
+3.58327 * Time
+0.25462 * Temperature
+2.22997 * Catalyst
+0.83491 * Time^2
+0.074772 * Temperature^2
+0.057094 * Catalyst^2
-0.38750 * Time * Temperature
-0.037500 * Time * Catalyst
+0.31250 * Temperature * Catalyst

Because many of the terms are insignificant, the reduced quadratic model is fit as follows:

Design Expert Output

Response: Activity
ANOVA for Response Surface Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	253.20	3	84.40	39.63	< 0.0001	significant
A	*175.35*	*1*	*175.35*	*82.34*	*< 0.0001*	
C	*67.91*	*1*	*67.91*	*31.89*	*< 0.0001*	
A^2	*9.94*	*1*	*9.94*	*4.67*	*0.0463*	
Residual	34.07	16	2.13			
Lack of Fit	*30.42*	*11*	*2.77*	*3.78*	*0.0766*	*not significant*
Pure Error	*3.65*	*5*	*0.73*			
Cor Total	287.28	19				

The Model F-value of 39.63 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	1.46	R-Squared	0.8814
Mean	60.51	Adj R-Squared	0.8591
C.V.	2.41	Pred R-Squared	0.6302
PRESS	106.24	Adeq Precision	20.447

	Coefficient		Standard	95% CI	95% CI	

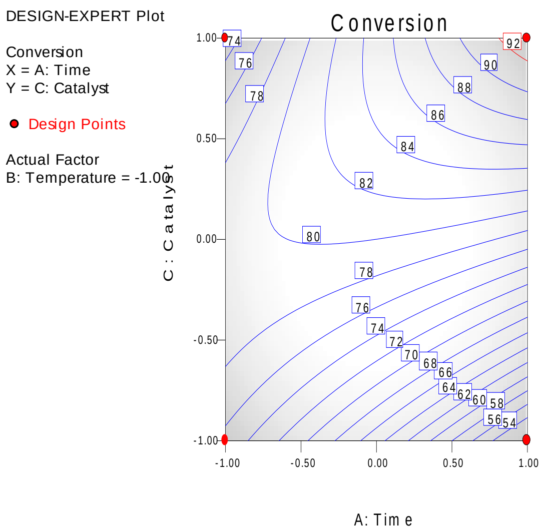
Factor	Estimate	DF	Error	Low	High	VIF
Intercept	59.95	1	0.42	59.06	60.83	
A-Time	3.58	1	0.39	2.75	4.42	1.00
C-Catalyst	2.23	1	0.39	1.39	3.07	1.00
A^2	0.82	1	0.38	0.015	1.63	1.00

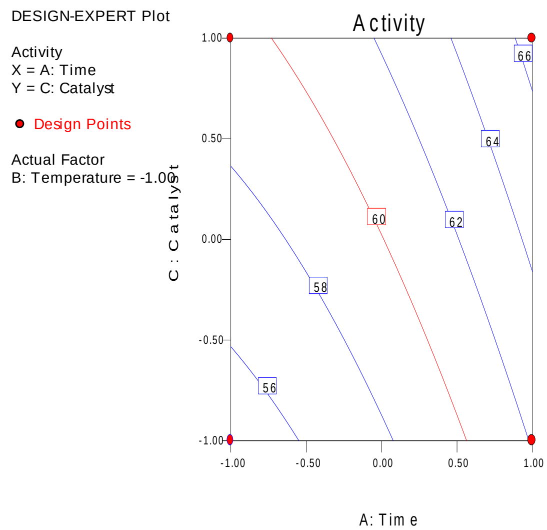
Final Equation in Terms of Coded Factors:

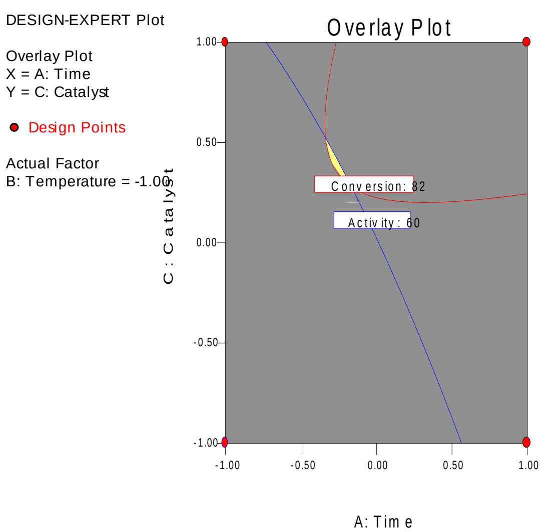
Activity =
+59.95
+3.58 * A
+2.23 * C
+0.82 * A2

Final Equation in Terms of Actual Factors:

Activity =
+59.94802
+3.58327 * Time
+2.22997 * Catalyst
+0.82300 * Time2







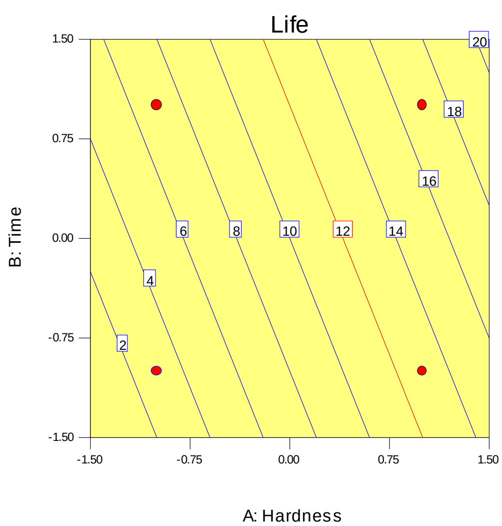
The contour plots visually describe the models while the overlay plots identifies the acceptable region for the process.

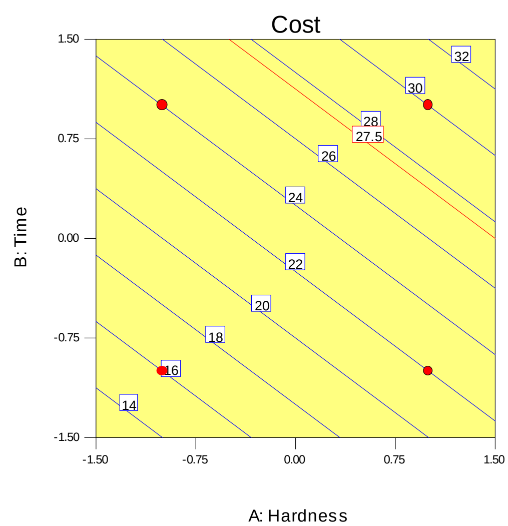
11.13. A manufacturer of cutting tools has developed two empirical equations for tool life in hours (y_1) and for tool cost in dollars (y_2). Both models are linear functions of steel hardness (x_1) and manufacturing time (x_2). The two equations are

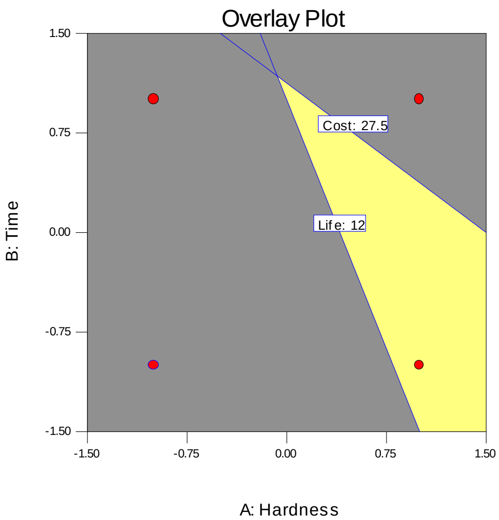
$$\hat{y}_1 = 10 + 5x_1 + 2x_2$$
$$\hat{y}_2 = 23 + 3x_1 + 4x_2$$

and both equations are valid over the range $-1.5 \le x_i \le 1.5$. Unit tool cost must be below \$27.50 and life must exceed 12 hours for the product to be competitive. Is there a feasible set of operating conditions for this process? Where would you recommend that the process be run?

The contour plots below graphically describe the two models. The overlay plot identifies the feasible operating region for the process.







$$10 + 5x_1 + 2x_2 \geq 12$$
$$23 + 3x_1 + 4x_2 \leq 27.50$$

11.14. A central composite design is run in a chemical vapor deposition process, resulting in the experimental data shown in Table P11.7. Four experimental units were processed simultaneously on each run of the design, and the responses are the mean and variance of thickness, computed across the four units.

Table P11.7

x_1	x_2	$\bar{y}$	s^2
-1	-1	360.6	6.689
1	-1	445.2	14.230
-1	1	412.1	7.088
1	1	601.7	8.586
1.414	0	518.0	13.130
-1.414	0	411.4	6.644
0	1.414	497.6	7.649
0	-1.414	397.6	11.740
0	0	530.6	7.836
0	0	495.4	9.306
0	0	510.2	7.956
0	0	487.3	9.127

(a) Fit a model to the mean response. Analyze the residuals.

Design Expert Output

Response: Mean Thick
ANOVA for Response Surface Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	47644.26	5	9528.85	16.12	0.0020	significant
A	*22573.36*	*1*	*22573.36*	*38.19*	*0.0008*	
B	*15261.91*	*1*	*15261.91*	*25.82*	*0.0023*	
A^2	*2795.58*	*1*	*2795.58*	*4.73*	*0.0726*	
B^2	*5550.74*	*1*	*5550.74*	*9.39*	*0.0221*	
AB	*2756.25*	*1*	*2756.25*	*4.66*	*0.0741*	
Residual	3546.83	6	591.14			
Lack of Fit	*2462.04*	*3*	*820.68*	*2.27*	*0.2592*	*not significant*

Pure Error	*1084.79*	*3*	*361.60*
Cor Total	51191.09	11	

The Model F-value of 16.12 implies the model is significant. There is only a 0.20% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	24.31	R-Squared	0.9307
Mean	472.31	Adj R-Squared	0.8730
C.V.	5.15	Pred R-Squared	0.6203
PRESS	19436.37	Adeq Precision	11.261

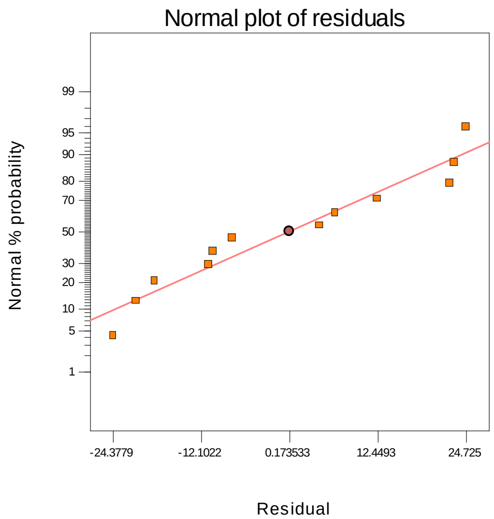
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	505.88	1	12.16	476.13	535.62	
A-x1	53.12	1	8.60	32.09	74.15	1.00
B-x2	43.68	1	8.60	22.64	64.71	1.00
A^2	-20.90	1	9.61	-44.42	2.62	1.04
B^2	-29.45	1	9.61	-52.97	-5.93	1.04
AB	26.25	1	12.16	-3.50	56.00	1.00

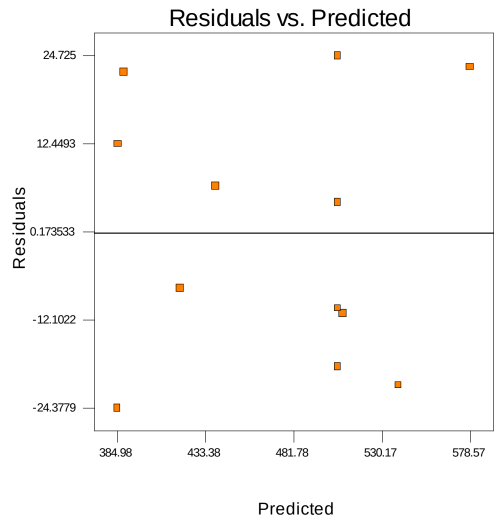
Final Equation in Terms of Coded Factors:

Mean Thick	=
+505.88	
+53.12	* A
+43.68	* B
-20.90	* A^2
-29.45	* B^2
+26.25	* A * B

Final Equation in Terms of Actual Factors:

Mean Thick	=
+505.87500	
+53.11940	* x1
+43.67767	* x2
-20.90000	* $x1^2$
-29.45000	* $x2^2$
+26.25000	* x1 * x2





A modest deviation from normality can be observed in the Normal Plot of Residuals; however, not enough to be concerned.

(b) Fit a model to the variance response. Analyze the residuals.

Design Expert Output

Response: Var Thick

ANOVA for Response Surface 2FI Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	65.80	3	21.93	35.86	< 0.0001	significant
A	*41.46*	*1*	*41.46*	*67.79*	*< 0.0001*	
B	*15.21*	*1*	*15.21*	*24.87*	*0.0011*	
AB	*9.13*	*1*	*9.13*	*14.93*	*0.0048*	
Residual	4.89	8	0.61			
Lack of Fit	*3.13*	*5*	*0.63*	*1.06*	*0.5137*	*not significant*
Pure Error	*1.77*	*3*	*0.59*			
Cor Total	70.69	11				

The Model F-value of 35.86 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.78	R-Squared	0.9308
Mean	9.17	Adj R-Squared	0.9048
C.V.	8.53	Pred R-Squared	0.8920
PRESS	7.64	Adeq Precision	18.572

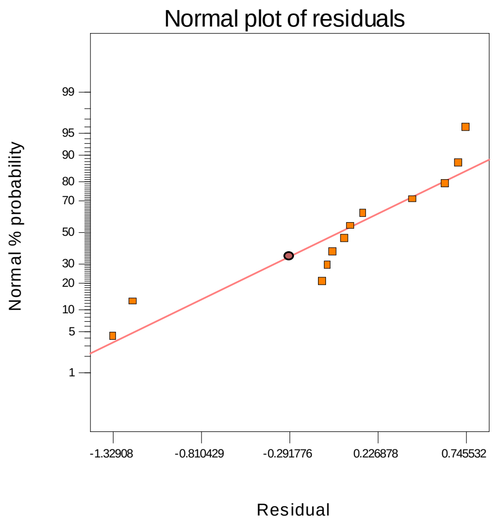
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	9.17	1	0.23	8.64	9.69	
A-x1	2.28	1	0.28	1.64	2.91	1.00
B-x2	-1.38	1	0.28	-2.02	-0.74	1.00
AB	-1.51	1	0.39	-2.41	-0.61	1.00

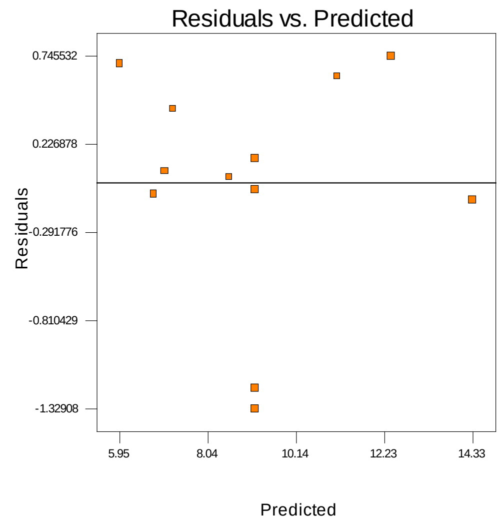
Final Equation in Terms of Coded Factors:

Var Thick =
+9.17
+2.28 * A
-1.38 * B
-1.51 * A * B

Final Equation in Terms of Actual Factors:

Var Thick =
+9.16508
+2.27645 * x1
-1.37882 * x2
-1.51075 * x1 * x2





The residual plots are not acceptable. A transformation should be considered. If not successful at correcting the residual plots, further investigation into the two apparently unusual points should be made.

(c) Fit a model to the $\ln(s^2)$. Is this model superior to the one you found in part (b)?

Design Expert Output

Response: Var Thick Transform: Natural log Constant: 0

ANOVA for Response Surface 2FI Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	0.67	3	0.22	36.94	< 0.0001	significant
A	*0.46*	*1*	*0.46*	*74.99*	*< 0.0001*	
B	*0.14*	*1*	*0.14*	*22.80*	*0.0014*	
AB	*0.079*	*1*	*0.079*	*13.04*	*0.0069*	
Residual	0.049	8	6.081E-003			
Lack of Fit	*0.024*	*5*	*4.887E-003*	*0.61*	*0.7093*	*not significant*
Pure Error	*0.024*	*3*	*8.071E-003*			
Cor Total	0.72	11				

The Model F-value of 36.94 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.078	R-Squared	0.9327
Mean	2.18	Adj R-Squared	0.9074
C.V.	3.57	Pred R-Squared	0.8797
PRESS	0.087	Adeq Precision	18.854

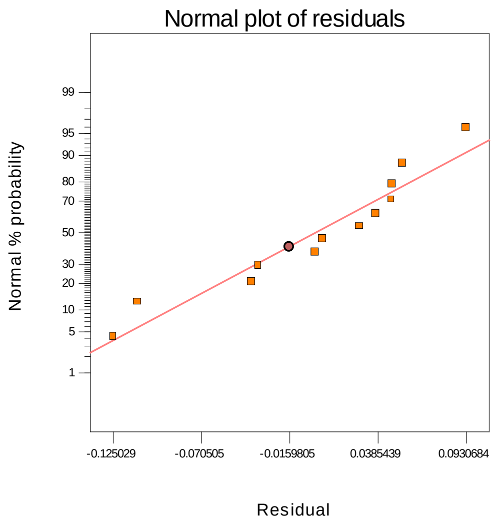
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	2.18	1	0.023	2.13	2.24	
A-x1	0.24	1	0.028	0.18	0.30	1.00
B-x2	-0.13	1	0.028	-0.20	-0.068	1.00
AB	-0.14	1	0.039	-0.23	-0.051	1.00

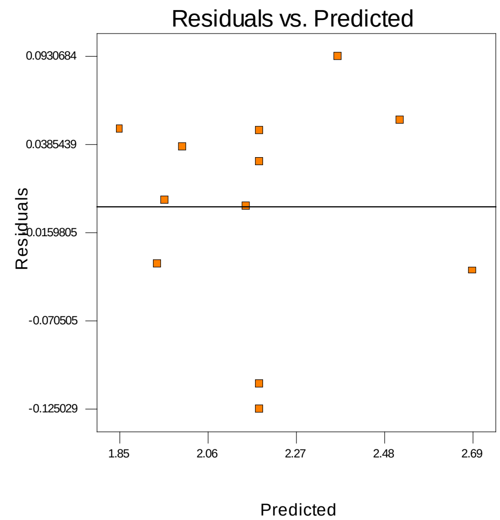
Final Equation in Terms of Coded Factors:

Ln(Var Thick) =
+2.18
+0.24 * A
-0.13 * B
-0.14 * A * B

Final Equation in Terms of Actual Factors:

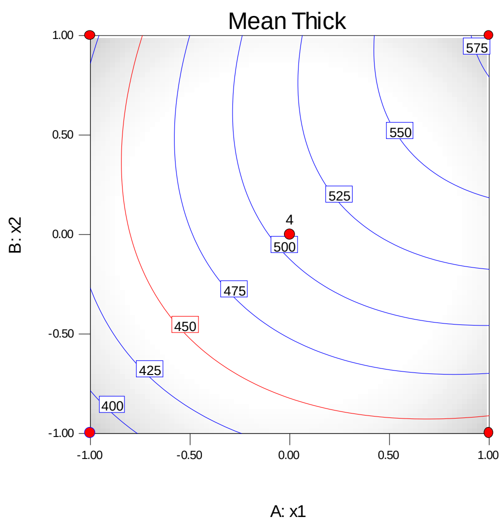
Ln(Var Thick)	=	
	+2.18376	
	+0.23874	* x1
	-0.13165	* x2
	-0.14079	* x1 * x2

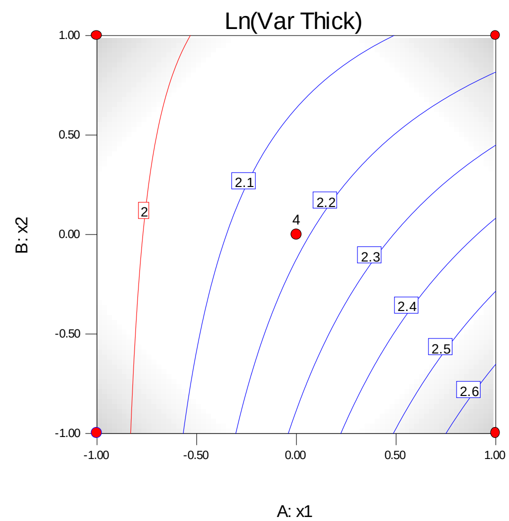


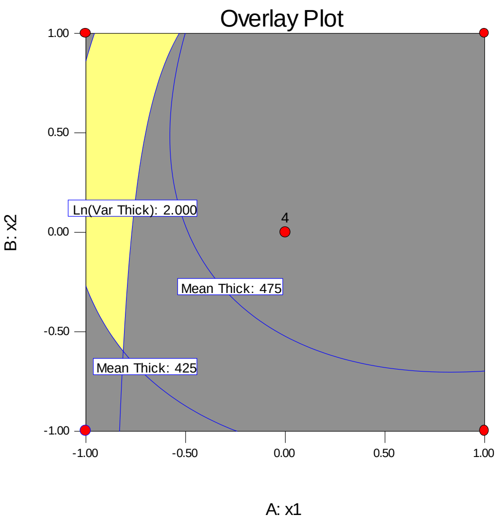


The residual plots are much improved following the natural log transformation; however, the two runs still appear to be somewhat unusual and should be investigated further. They will be retained in the analysis.

(d) Suppose you want the mean thickness to be in the interval 450±25. Find a set of operating conditions that achieve the objective and simultaneously minimize the variance.







The contour plots describe the two models while the overlay plot identifies the acceptable region for the process.

(e) Discuss the variance minimization aspects of part (d). Have you minimized total process variance?

The within run variance has been minimized; however, the run-to-run variation has not been minimized in the analysis. This may not be the most robust operating conditions for the process.

11.15. Verify that an orthogonal first-order design is also first-order rotatable.

To show that a first order orthogonal design is also first order rotatable, consider

$$V(\hat{y}) = V(\hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i x_i) = V(\hat{\beta}_0) + \sum_{i=1}^{k} x_i^2 V(\hat{\beta}_i)$$

since all covariances between $\hat{\beta}_i$ and $\hat{\beta}_j$ are zero, due to design orthogonality. Furthermore, we have:

$$V(\hat{\beta}_0) = V(\hat{\beta}_1) = V(\hat{\beta}_2) = \ldots = V(\hat{\beta}_k) = \frac{\sigma^2}{n}\text{, so}$$

$$V(\hat{y}) = \frac{\sigma^2}{n} + \frac{\sigma^2}{n} \sum_{i=1}^{k} x_i^2$$

$$V(\hat{y}) = \frac{\sigma^2}{n} \left(1 + \sum_{i=1}^{k} x_i^2 \right)$$

which is a function of distance from the design center (i.e. **x**=**0**), and not direction. Thus the design must be rotatable. Note that *n* is, in general, the number of points in the exterior portion of the design. If there are n_c center points, then $V(\hat{\beta}_0) = \frac{\sigma^2}{(n + n_c)}$.

11.16. Show that augmenting a 2^k design with n_C center points does not affect the estimates of the β_i (i=1, 2, . . . , k), but that the estimate of the intercept β_0 is the average of all $2^k + n_C$ observations.

In general, the **X** matrix for the 2^k design with n_c center points and the **y** vector would be:

$$\begin{array}{cc} & \begin{array}{ccccc} \beta_0 & \beta_1 & \beta_2 & \dots & \beta_k \end{array} \\ \mathbf{X} = & \left[\begin{array}{ccccc} 1 & -1 & -1 & \square & -1 \\ 1 & 1 & -1 & \square & -1 \\ \square & \square & \square & \square & \square \\ 1 & 1 & 1 & \square & 1 \\ \hline 1 & 0 & 0 & \square & 0 \\ 1 & 0 & 0 & \square & 0 \\ \square & \square & \square & \square & \square \\ 1 & 0 & 0 & \square & 0 \end{array}\right] \end{array}$$

← The upper half of the matrix is the usual ± 1 notation of the 2^k design

← The lower half of the matrix represents the center points (n_c rows)

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \square \\ y_{2^k} \\ --- \\ n_{0_1} \\ n_{0_2} \\ \square \\ n_{0_c} \end{bmatrix}$$

← 2^k+n_c observations
.

$$\mathbf{X'X} = \begin{bmatrix} 2^k + n_c & 0 & \square & 0 \\ & 2^k & \square & 0 \\ & & \square & \square \\ & & & 2^k \end{bmatrix} \qquad \mathbf{X'y} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \square \\ g_k \end{bmatrix}$$

← Grand total of all 2^k+n_c observations

← usual contrasts from 2^k

Therefore, $\ddot{\beta}_0 = \dfrac{g_0}{2^k + n_c}$, which is the average of all $\left(2^k + n_c\right)$observations, while $\ddot{\beta}_i = \dfrac{g_i}{2^k}$, which does not depend on the number of center points, since in computing the contrasts g_i, all observations at the center are multiplied by zero.

11.17. ***The rotatable central composite design.*** It can be shown that a second-order design is rotatable if $\sum_{u=1}^n x_{iu}^a x_{ju}^b = 0$ if a or b (or both) are odd and if $\sum_{u=1}^n x_{iu}^4 = 3\sum_{u=1}^n x_{iu}^2 x_{ju}^2$. Show that for the central composite design these conditions lead to $\alpha = \left(n_f\right)^{/4}$ for rotatability, where n_f is the number of points in the factorial portion.

The balance between +1 and -1 in the factorial columns and the orthogonality among certain columns in the X matrix for the central composite design will result in all odd moments being zero. To solve for α use the following relations:

$$\sum_{u=1}^n x_{iu}^4 = n_f + 2\alpha^4, \qquad \sum_{u=1}^n x_{iu}^2 x_{ju}^2 = n_f$$

then

$$\sum_{u=1}^{n} x_{iu}^4 = 3\sum_{u=1}^{n} x_{iu}^2 x_{ju}^2$$
$$n_f + 2\alpha^4 = 3(n_f)$$
$$2\alpha^4 = 2n_f$$
$$\alpha^4 = n_f$$
$$\alpha = \sqrt[4]{n_f}$$

11.18. Verify that the central composite design shown in Table P11.8 blocks orthogonally.

Table P11.8

Block 1			Block 2			Block 3		
x_1	x_2	x_3	x_1	x_2	x_3	x_1	x_2	x_3
0	0	0	0	0	0	-1.633	0	0
0	0	0	0	0	0	1.633	0	0
1	1	1	1	1	-1	0	-1.633	0
1	-1	-1	1	-1	1	0	1.633	0
-1	-1	1	-1	1	1	0	0	-1.633
-1	1	-1	-1	-1	-1	0	0	1.633
						0	0	0
						0	0	0

Note that each block is an orthogonal first order design, since the cross products of elements in different columns add to zero for each block. To verify the second condition, choose a column, say column x_2. Now

$$\sum_{u=1}^{k} x_{2u}^2 = 13.334 \text{ , and } n=20$$

For blocks 1 and 2,

$$\sum_{m} x_{2m}^2 = 4 \text{ , } n_{\mathrm{m}}=6$$

So

$$\frac{\sum_{m} x_{2m}^2}{\sum_{u=1}^{n} x_{2u}^2} = \frac{n_m}{n}$$

$$\frac{4}{13.334} = \frac{6}{20}$$

$$0.3 = 0.3$$

and condition 2 is satisfied by blocks 1 and 2. For block 3, we have

$$\sum_m x_{2m}^2 = 5.334 \text{ , } n_m = 8 \text{, so}$$

$$\frac{\sum_m x_{2m}^2}{\sum_{u=1}^{n} x_{2u}^2} = \frac{n_m}{n}$$

$$\frac{5.334}{13.334} = \frac{8}{20}$$

$$0.4 = 0.4$$

And condition 2 is satisfied by block 3. Similar results hold for the other columns.

11.19. *Blocking in the central composite design*. Consider a central composite design for k = 4 variables in two blocks. Can a rotatable design always be found that blocks orthogonally?

To run a central composite design in two blocks, assign the n_f factorial points and the n_{c1} center points to block 1 and the 2k axial points plus n_{c2} center points to block 2. Both blocks will be orthogonal first order designs, so the first condition for orthogonal blocking is satisfied.

The second condition implies that

$$\frac{\sum_m x_{im}^2 (block1)}{\sum_m x_{im}^2 (block2)} = \frac{n_f + n_{c1}}{2k + n_{c2}}$$

However, $\sum_m x_{im}^2 = n_f$ in block 1 and $\sum_m x_{im}^2 = 2\alpha^2$ in block 2, so

$$\frac{n_f}{2\alpha^2} = \frac{n_f + n_{c1}}{2k + n_{c2}}$$

Which gives:

$$\alpha = \left[\frac{n_f (2k + n_{c2})}{2(n_f + n_{c1})}\right]^{\frac{1}{2}}$$

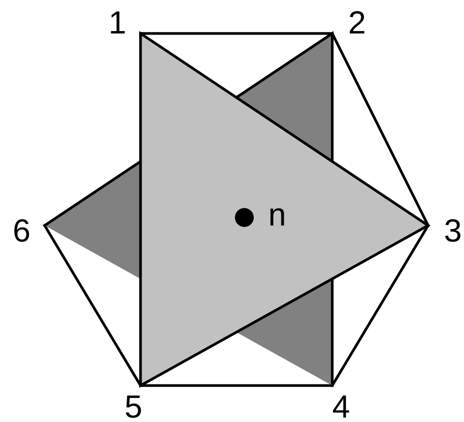
Since $\alpha = \sqrt[4]{n_f}$ if the design is to be rotatable, then the design must satisfy

$$n_f = \left[\frac{n_f (2k + n_{c2})}{2(n_f + n_{c1})}\right]^{2}$$

It is not possible to find rotatable central composite designs which block orthogonally for all k. For example, if k=3, the above condition cannot be satisfied. For k=2, there must be an equal number of center points in each block, i.e. $n_{c1} = n_{c2}$. For k=4, we must have $n_{c1} = 4$ and $n_{c2} = 2$.

11.20. How could a hexagon design be run in two orthogonal blocks?

The hexagonal design can be blocked as shown below. There are $n_{c1} = n_{c2} = n_c$ center points with n_c even.



Put the points 1,3,and 5 in block 1 and 2,4,and 6 in block 2. Note that each block is a simplex.

11.21. Yield during the first four cycles of a chemical process is shown in the following table. The variables are percent concentration (x_1) at levels 30, 31, and 32 and temperature (x_2) at 140, 142, and 144°F. Analyze by EVOP methods.

	Conditions				
Cycle	(1)	(2)	(3)	(4)	(5)
1	60.7	59.8	60.2	64.2	57.5
2	59.1	62.8	62.5	64.6	58.3
3	56.6	59.1	59.0	62.3	61.1
4	60.5	59.8	64.5	61.0	60.1

Cycle: n=1 Phase 1

Calculation of Averages							Calculation of Standard Deviation
Operating Conditions		(1)	(2)	(3)	(4)	(5)	
(i)	Previous Cycle Sum						Previous Sum S=
(ii)	Previous Cycle Average						Previous Average =
(iii)	New Observation	60.7	59.8	60.2	64.2	57.5	New S=Range x $f_{k,n}$
(iv)	Differences						Range=
(v)	New Sums	60.7	59.8	60.2	64.2	57.5	New Sum S=
(vi)	New Averages	60.7	59.8	60.2	64.2	57.5	New average S = New Sum S/(n-1)=

Calculation of Effects		Calculation of Error Limits
$A = \frac{1}{2}\left(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5\right) =$	3.55	For New Average: $\left(\frac{2}{\sqrt{n}}\right)S =$
$B = \frac{1}{2}\left(\bar{y}_3 - \bar{y}_4 - \bar{y}_2 + \bar{y}_5\right) =$	-3.15	For New Effects: $\left(\frac{2}{\sqrt{n}}\right)S =$

$$AB = \frac{1}{2}\left(\bar{y}_3 - \bar{y}_4 + \bar{y}_2 - \bar{y}_5\right) = \quad -0.85 \qquad \text{For CIM: } \left(\frac{1.78}{\sqrt{n}}\right)S =$$

$$CIM = \frac{1}{5}\left(\bar{y}_3 + \bar{y}_4 + \bar{y}_2 + \bar{y}_5 - 4\bar{y}_1\right) = \quad -0.22$$

Cycle: n=2 Phase 1

Calculation of Averages							Calculation of Standard Deviation
Operating Conditions		(1)	(2)	(3)	(4)	(5)	
(i)	Previous Cycle Sum	60.7	59.8	60.2	64.2	57.5	Previous Sum S=
(ii)	Previous Cycle Average	60.7	59.8	60.2	64.2	57.5	Previous Average =
(iii)	New Observation	59.1	62.8	62.5	64.6	58.3	New S=Range x $f_{k,n}$=1.38
(iv)	Differences	1.6	-3.0	-2.3	-0.4	-0.8	Range=4.6
(v)	New Sums	119.8	122.6	122.7	128.8	115.8	New Sum S=1.38
(vi)	New Averages	59.90	61.30	61.35	64.40	57.90	New average S = New Sum S/(n-1)=1.38

Calculation of Effects — Calculation of Error Limits

$$A = \frac{1}{2}\left(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5\right) = \quad 3.28 \qquad \text{For New Average: } \left(\frac{2}{\sqrt{n}}\right)S = \quad 1.95$$

$$B = \frac{1}{2}\left(\bar{y}_3 - \bar{y}_4 - \bar{y}_2 + \bar{y}_5\right) = \quad -3.23 \qquad \text{For New Effects: } \left(\frac{2}{\sqrt{n}}\right)S = \quad 1.95$$

$$AB = \frac{1}{2}\left(\bar{y}_3 - \bar{y}_4 + \bar{y}_2 - \bar{y}_5\right) = \quad 0.18 \qquad \text{For CIM: } \left(\frac{1.78}{\sqrt{n}}\right)S = \quad 1.74$$

$$CIM = \frac{1}{5}\left(\bar{y}_3 + \bar{y}_4 + \bar{y}_2 + \bar{y}_5 - 4\bar{y}_1\right) = \quad 1.07$$

Cycle: n=3 Phase 1

Calculation of Averages							Calculation of Standard Deviation
Operating Conditions		(1)	(2)	(3)	(4)	(5)	
(i)	Previous Cycle Sum	119.8	122.6	122.7	128.8	115.8	Previous Sum S=1.38
(ii)	Previous Cycle Average	59.90	61.30	61.35	64.40	57.90	Previous Average =1.38
(iii)	New Observation	56.6	59.1	59.0	62.3	61.1	New S=Range x $f_{k,n}$=2.28
(iv)	Differences	3.30	2.20	2.35	2.10	-3.20	Range=6.5
(v)	New Sums	176.4	181.7	181.7	191.1	176.9	New Sum S=3.66
(vi)	New Averages	58.80	60.57	60.57	63.70	58.97	New average S = New Sum S/(n-1)=1.83

Calculation of Effects — Calculation of Error Limits

$$A = \frac{1}{2}\left(\bar{y}_3 + \bar{y}_4 - \bar{y}_2 - \bar{y}_5\right) = \quad 2.37 \qquad \text{For New Average: } \left(\frac{2}{\sqrt{n}}\right)S = \quad 2.11$$

$$B = \frac{1}{2}\left(\bar{y}_3 - \bar{y}_4 - \bar{y}_2 + \bar{y}_5\right) = \quad -2.37 \qquad \text{For New Effects: } \left(\frac{2}{\sqrt{n}}\right)S = \quad 2.11$$

$$AB = \frac{1}{2}\left(\bar{y}_3 - \bar{y}_4 + \bar{y}_2 - \bar{y}_5\right) = \quad -0.77 \qquad \text{For CIM: } \left(\frac{1.78}{\sqrt{n}}\right)S = \quad 1.88$$

$$CIM = \frac{1}{5}\left(\bar{y}_3 + \bar{y}_4 + \bar{y}_2 + \bar{y}_5 - 4\bar{y}_1\right) = \quad 1.72$$

Cycle: n=4 Phase 1

Calculation of Averages							Calculation of Standard Deviation
Operating Conditions		(1)	(2)	(3)	(4)	(5)	
(i)	Previous Cycle Sum	176.4	181.7	181.7	191.1	176.9	Previous Sum S=3.66

(ii)	Previous Cycle Average	58.80	60.57	60.57	63.70	58.97	Previous Average =1.83
(iii)	New Observation	60.5	59.8	64.5	61.0	60.1	New S=Range x $f_{k,n}$=2.45
(iv)	Differences	-1.70	0.77	-3.93	2.70	-1.13	Range=6.63
(v)	New Sums	236.9	241.5	246.2	252.1	237.0	New Sum S=6.11
(vi)	New Averages	59.23	60.38	61.55	63.03	59.25	New average S = New Sum S/(n-1)=2.04

Calculation of Effects		Calculation of Error Limits	
$A=\frac{1}{2}\left(\bar{y}_3+\bar{y}_4-\bar{y}_2-\bar{y}_5\right)=$	2.48	For New Average: $\left(\frac{2}{\sqrt{n}}\right)S=$	2.04
$B=\frac{1}{2}\left(\bar{y}_3-\bar{y}_4-\bar{y}_2+\bar{y}_5\right)=$	-1.31	For New Effects: $\left(\frac{2}{\sqrt{n}}\right)S=$	2.04
$AB=\frac{1}{2}\left(\bar{y}_3-\bar{y}_4+\bar{y}_2-\bar{y}_5\right)=$	-0.18	For CIM: $\left(\frac{1.78}{\sqrt{n}}\right)S=$	1.82
$CIM=\frac{1}{5}\left(\bar{y}_3+\bar{y}_4+\bar{y}_2+\bar{y}_5-4\bar{y}_1\right)=$	1.46		

From studying cycles 3 and 4, it is apparent that *A* (and possibly *B*) has a significant effect. A new phase should be started following cycle 3 or 4.

11.22. Suppose that we approximate a response surface with a model of order d_1, such as $\mathbf{y}=\mathbf{X}_1\boldsymbol{\beta}_1+\boldsymbol{\varepsilon}$, when the true surface is described by a model of order $d_2>d_1$; that is $E(\mathbf{y}) = \mathbf{X}_1\boldsymbol{\beta}_1+\mathbf{X}_2\boldsymbol{\beta}_2$.

(a) Show that the regression coefficients are biased, that is, that $E(\hat{\boldsymbol{\beta}}_1)=\boldsymbol{\beta}_1+\mathbf{A}\boldsymbol{\beta}_2$, where $\mathbf{A}=(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$. **A** is usually called the alias matrix.

$$
\begin{aligned}
E\left[\hat{\boldsymbol{\beta}}_1\right] &= E\left[\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{y}\right] \\
&= \left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'E[\mathbf{y}] \\
&= \left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\left(\mathbf{X}_1\boldsymbol{\beta}_1+\mathbf{X}_2\boldsymbol{\beta}_2\right) \\
&= \left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{X}_1\boldsymbol{\beta}_1+\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{X}_2\boldsymbol{\beta}_2 \\
&= \boldsymbol{\beta}_1+\mathbf{A}\boldsymbol{\beta}_2
\end{aligned}
$$

where $\mathbf{A}=\left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{X}_2$

(b) If d_1=1 and d_2=2, and a full 2^k is used to fit the model, use the result in part (a) to determine the alias structure.

In this situation, we have assumed the true surface to be first order, when it is really second order. If a full factorial is used for *k*=2, then

$$
\mathbf{X}_1=\begin{array}{c} \begin{array}{ccc}\beta_0 & \beta_1 & \beta_2\end{array} \\ \begin{bmatrix} 1 & -1 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}\end{array}
\quad
\mathbf{X}_2=\begin{array}{c} \begin{array}{ccc}\beta_{11} & \beta_{22} & \beta_{12}\end{array} \\ \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}\end{array}
\quad \text{and} \quad
\mathbf{A}=\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}
$$

$$\text{Then, } \mathbf{E}\left[\hat{\boldsymbol{\beta}}_1\right] = \mathbf{E}\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix} = \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix} + \begin{bmatrix}1 & 1 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\beta_{11}\\ \beta_{22}\\ \beta_{12}\end{bmatrix} = \begin{bmatrix}\beta_0+\beta_{11}+\beta_{22}\\ \beta_1\\ \beta_2\end{bmatrix}$$

The pure quadratic terms bias the intercept.

(c) If d_1=1, d_2=2 and k=3, find the alias structure assuming that a 2^{3-1} design is used to fit the model.

$$\mathbf{X}_1 = \begin{array}{c}\begin{array}{cccc}\beta_0 & \beta_1 & \beta_2 & \beta_3\end{array}\\ \begin{bmatrix}1 & -1 & -1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 1 & 1 & 1 & 1\end{bmatrix}\end{array} \quad \mathbf{X}_2 = \begin{array}{c}\begin{array}{cccccc}\beta_{11} & \beta_{22} & \beta_{33} & \beta_{12} & \beta_{13} & \beta_{23}\end{array}\\ \begin{bmatrix}1 & 1 & 1 & 1 & -1 & -1\\ 1 & 1 & 1 & -1 & -1 & 1\\ 1 & 1 & 1 & -1 & 1 & -1\\ 1 & 1 & 1 & 1 & 1 & 1\end{bmatrix}\end{array} \quad \text{and} \quad \mathbf{A} = \begin{bmatrix}1 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\end{bmatrix}$$

$$\text{Then, } \mathbf{E}\left[\hat{\boldsymbol{\beta}}_1\right] = \mathbf{E}\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\end{bmatrix} = \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} + \begin{bmatrix}1 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}\beta_{11}\\ \beta_{22}\\ \beta_{33}\\ \beta_{12}\\ \beta_{13}\\ \beta_{23}\end{bmatrix} = \begin{bmatrix}\beta_0+\beta_{11}+\beta_{22}+\beta_{33}\\ \beta_1+\beta_{23}\\ \beta_2+\beta_{13}\\ \beta_3+\beta_{12}\end{bmatrix}$$

(d) If d_1=1, d_2=2, k=3, and the simplex design in Problem 11.3 is used to fit the model, determine the alias structure and compare the results with part (c).

$$\mathbf{X}_1 = \begin{array}{c}\begin{array}{cccc}\beta_0 & \beta_1 & \beta_2 & \beta_3\end{array}\\ \begin{bmatrix}0 & 0 & \sqrt{2} & -1\\ 2 & -\sqrt{2} & 0 & 1\\ 0 & 0 & -\sqrt{2} & -1\\ 2 & -\sqrt{2} & 0 & 1\end{bmatrix}\end{array} \quad \mathbf{X}_2 = \begin{array}{c}\begin{array}{cccccc}\beta_{11} & \beta_{22} & \beta_{33} & \beta_{12} & \beta_{13} & \beta_{23}\end{array}\\ \begin{bmatrix}0 & 2 & 1 & 0 & 0 & -\sqrt{2}\\ 2 & 0 & 1 & 0 & -\sqrt{2} & 0\\ 0 & 2 & 1 & 0 & 0 & -\sqrt{2}\\ 2 & 0 & 1 & 0 & -\sqrt{2} & 0\end{bmatrix}\end{array} \quad \text{and } \mathbf{A} = \begin{bmatrix}1 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & -1\\ 1 & -1 & 0 & 0 & 0 & 0\end{bmatrix}$$

$$\text{Then, } \mathbf{E}\left[\hat{\boldsymbol{\beta}}_1\right] = \mathbf{E}\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\end{bmatrix} = \begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\end{bmatrix} + \begin{bmatrix}1 & 1 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & -1\\ 1 & -1 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\beta_{11}\\ \beta_{22}\\ \beta_{33}\\ \beta_{12}\\ \beta_{13}\\ \beta_{23}\end{bmatrix} = \begin{bmatrix}\beta_0+\beta_{11}+\beta_{22}+\beta_{33}\\ \beta_1+\beta_{13}\\ \beta_2-\beta_{23}\\ \beta_3+\beta_{11}-\beta_{22}\end{bmatrix}$$

Notice that the alias structure is different from that found in the previous part for the 2^{3-1} design. In general, the **A** matrix will depend on which simplex design is used.

11.23. Suppose that you need to design an experiment to fit a quadratic model over the region $-1 \le x_i \le +1$, i=1,2 subject to the constraint $x_1 + x_2 \le 1$. If the constraint is violated, the process will not work properly. You can afford to make no more than n=12 runs. Set up the following designs:

(a) An "inscribed" CCD with center points at $x_1 = x_2 = 0$

x_1	x_2
-0.5	-0.5
0.5	-0.5
-0.5	0.5
0.5	0.5
-0.707	0
0.707	0
0	-0.707
0	0.707
0	0
0	0
0	0
0	0

(a) An "inscribed" CCD with center points at $x_1 = x_2 = -0.25$ so that a larger design could be fit within the constrained region

x_1	x_2
-1	-1
0.5	-1
-1	0.5
0.5	0.5
-1.664	-0.25
1.164	-0.25
-0.25	-1.664
-0.25	1.164
-0.25	-0.25
-0.25	-0.25
-0.25	-0.25
-0.25	-0.25

(b) An "inscribed" 3^2 factorial with center points at $x_1 = x_2 - 0.25$

x_1	x_2
-1	-1
-0.25	-1
0.5	-1
-1	-0.25
-0.25	-0.25
0.5	-0.25
-1	0.5
-0.25	0.5
0.5	0.5
-0.25	-0.25
-0.25	-0.25
-0.25	-0.25

(c) A D-optimal design.

x_1	x_2
-1	-1
1	-1
-1	1
1	0
0	1
0	0
-1	0
0	-1
0.5	0.5
-1	-1
1	-1
-1	1

(d) A modified D-optimal design that is identical to the one in part (c), but with all replicate runs at the design center.

x_1	x_2
1	0
0	0
0	1
-1	-1
1	-1
-1	1
-1	0
0	-1
0.5	0.5
0	0
0	0
0	0

(e) Evaluate the $\left|(\mathbf{X}'\mathbf{X})^{-1}\right|$ criteria for each design.

	(a)	(a)*	(b)	(c)	(d)
$\left|(\mathbf{X}'\mathbf{X})^{-1}\right|$	0.5	0.00005248	0.007217	0.0001016	0.0002294

(f) Evaluate the D-efficiency for each design relative to the D-optimal design in part (c).

	(a)	(a)*	(b)	(c)	(d)
D-efficiency	24.25%	111.64%	49.14%	100.00%	87.31%

(g) Which design would you prefer? Why?

The offset CCD, (a)*, is the preferred design based on the D-efficiency. Not only is it better than the D-optimal design, (c), but it maintains the desirable design features of the CCD.

11.24. Consider a 2^3 design for fitting a first-order model.

(a) Evaluate the D-criterion $\left|(\mathbf{X'X})^{-1}\right|$ for this design.

$$\left|(\mathbf{X'X})^{-1}\right| = 2.441\text{E-4}$$

(b) Evaluate the A-criterion $tr(\mathbf{X'X})^{-1}$ for this design.

$$tr(\mathbf{X'X})^{-1} = 0.5$$

(c) Find the maximum scaled prediction variance for this design. Is this design G-optimal?

$$v(\mathbf{x}) = \frac{NVar(\hat{y}(\mathbf{x}))}{\sigma^2} = N\mathbf{x}'^{(1)}(\mathbf{X'X})^{-1}\mathbf{x}^{(1)} = 4$$. Yes, this is a G-optimal design.

11.25. Repeat Problem 11.24 using a first order model with the two-factor interaction.

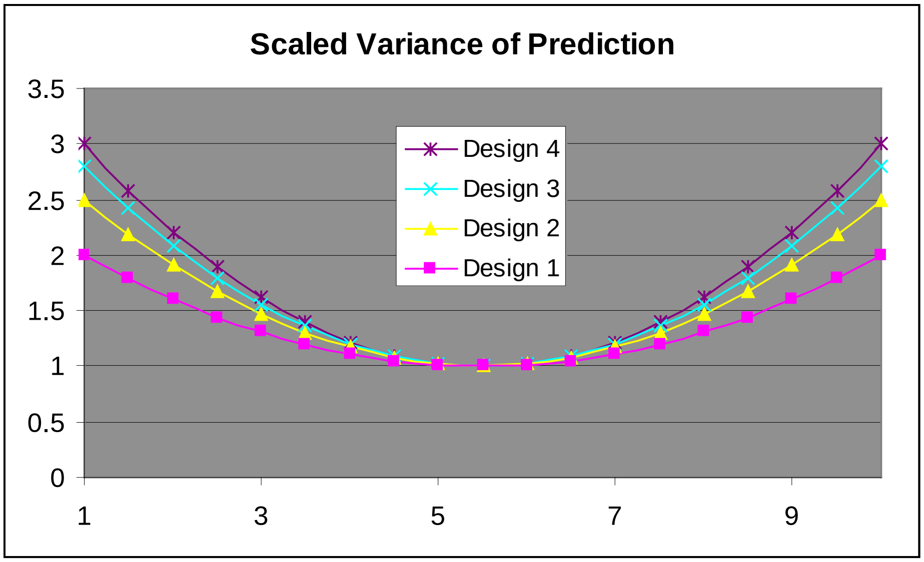
$$\left|(\mathbf{X'X})^{-1}\right| = 4.768\text{E-7}$$

$$tr(\mathbf{X'X})^{-1} = 0.875$$

$$v(\mathbf{x}) = \frac{NVar(\hat{y}(\mathbf{x}))}{\sigma^2} = N\mathbf{x}'^{(1)}(\mathbf{X'X})^{-1}\mathbf{x}^{(1)} = 7$$. Yes, this is a G-optimal design.

11.26. A chemical engineer wishes to fit a calibration curve for a new procedure used to measure the concentration of a particular ingredient in a product manufactured in his facility. Twelve samples can be prepared, having known concentration. The engineer wants to build a model for the measured concentrations. He suspects that a linear calibration curve will be adequate to model the measured concentration as a function of the known concentrations; that is, $y = \beta_0 + \beta_0 x + \varepsilon$, where x is the actual concentration. Four experimental designs are under consideration. Design 1 consists of 6 runs at known concentration 1 and 6 runs at known concentration 10. Design 2 consists of 4 runs at concentrations 1, 5.5, and 10. Design 3 consists of 3 runs at concentrations 1, 4, 7, and 10. Finally, design 4 consists of 3 runs at concentrations 1 and 10 and 6 runs at concentration 5.5.

(a) Plot the scaled variance of prediction for all four designs on the same graph over the concentration range. Which design would be preferable, in your opinion?



Because it has the lowest scaled variance of prediction at all points in the design space with the exception of 5.5, Design 1 is preferred.

(b) For each design calculate the determinant of $(\mathbf{X}'\mathbf{X})^{-1}$. Which design would be preferred according to the "D" criterion?

Design	$\left\|(\mathbf{X}'\mathbf{X})^{-1}\right\|$
1	0.000343
2	0.000514
3	0.000617
4	0.000686

Design 1 would be preferred.

(c) Calculate the D-efficiency of each design relative to the "best" design that you found in part b.

Design	D-efficiency
1	100.00%
2	81.65%
3	74.55%
4	70.71%

(d) For each design, calculate the average variance of prediction over the set of points given by x = 1, 1.5, 2, 2.5, . . ., 10. Which design would you prefer according to the V-criterion?

Average Variance of Prediction		
Design	Actual	Coded
1	1.3704	0.1142
2	1.5556	0.1296
3	1.6664	0.1389
4	1.7407	0.1451

Design 1 is still preferred based on the V-criterion.

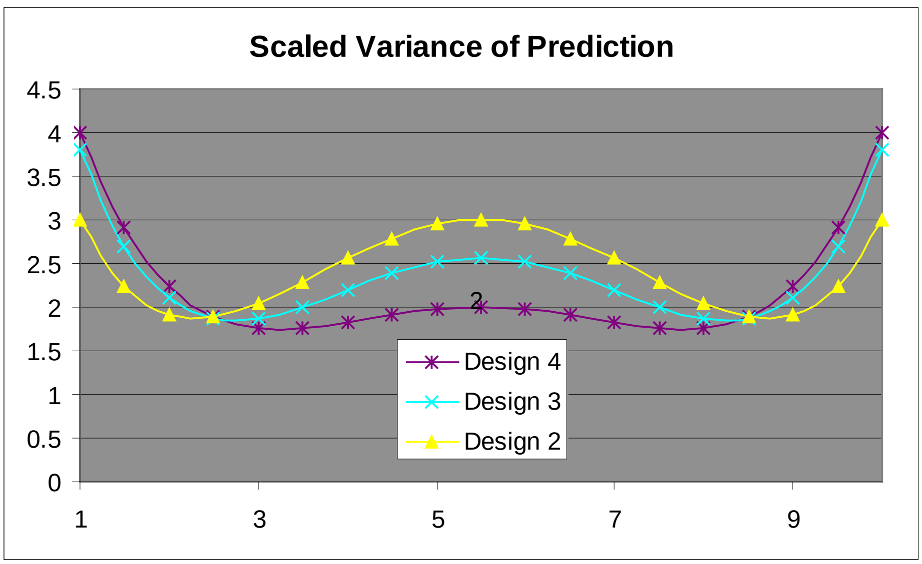
(e) Calculate the V-efficiency of each design relative to the best design you found in part (d).

Design	V-efficiency
1	100.00%
2	88.10%
3	82.24%
4	78.72%

(f) What is the G-efficiency of each design?

Design	G-efficiency
1	100.00%
2	80.00%
3	71.40%
4	66.70%

11.27. Rework Problem 11.26 assuming that the model the engineer wishes to fit is a quadratic. Obviously, only designs 2, 3, and 4 can now be considered.



Based on the plot, the preferred design would depend on the region of interest. Design 4 would be preferred if the center of the region was of interest; otherwise, Design 2 would be preferred.

Design	$\left|(\mathbf{X}'\mathbf{X})^{-1}\right|$
2	4.704E-07
3	6.351E-07
4	5.575E-07

Design 2 is preferred based on $\left|(\mathbf{X}'\mathbf{X})^{-1}\right|$.

Design	D-efficiency
2	100.00%
3	90.46%
4	94.49%

Average Variance of Prediction		
Design	Actual	Coded
2	2.441	0.2034
3	2.393	0.1994
4	2.242	0.1869

Design 4 is preferred.

Design	V-efficiency
2	91.89%
3	93.74%
4	100.00%

Design	G-efficiency
2	100.00%
3	79.00%
4	75.00%

11.28. Suppose that you want to fit a *second*-order model in $k = 5$ factors. You cannot afford more than 25 runs. Construct both a *D*-optimal and an *I*-optimal design for this situation. Compare the prediction variance properties of the designs. Which design would you prefer?

The following JMP outputs identify both *D*-optimal and *I*-optimal designs. The prediction variance profile shown in each JMP output is appreciably less for the *I*-optimal design as well as flatter in the middle of the design.

JMP Output

D-optimal Design

Factors

Add N Factors
1

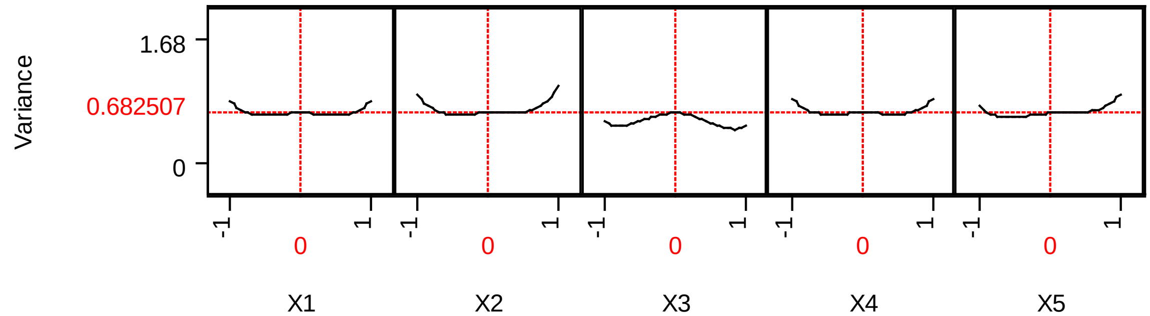
X1 Continuous
X2 Continuous
X3 Continuous
X4 Continuous
X5 Continuous

Design

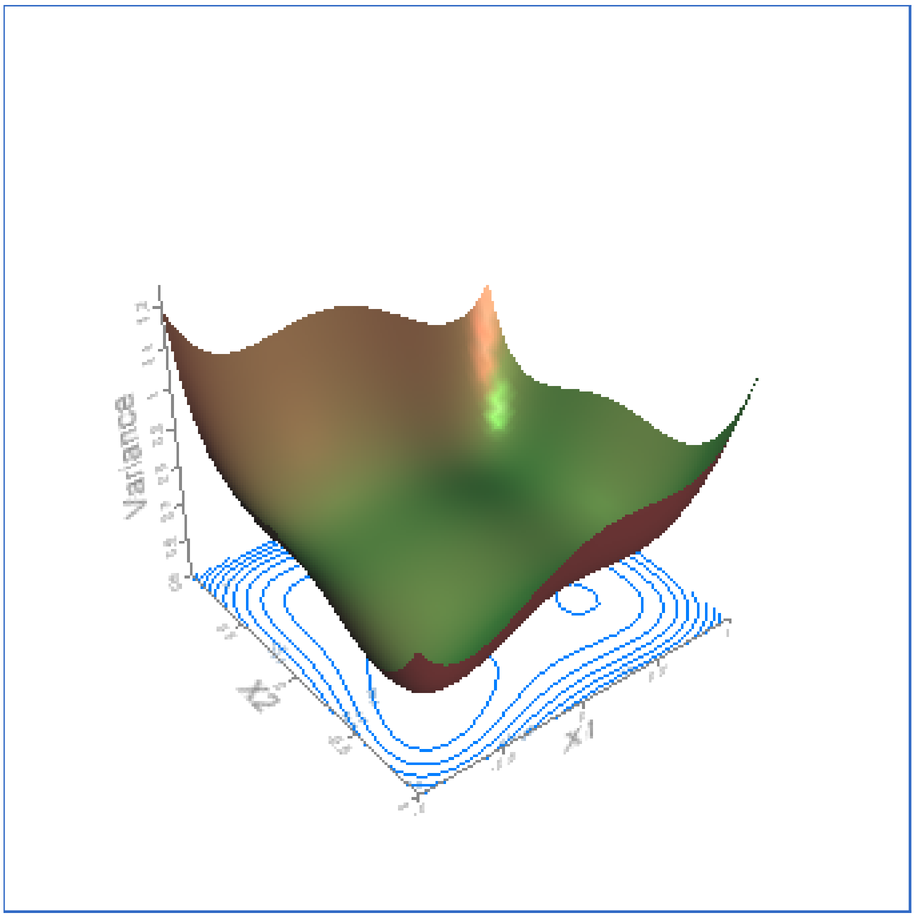
Run	X1	X2	X3	X4	X5
1	-1	-1	1	1	0
2	-1	1	1	-1	-1
3	-1	1	-1	-1	1
4	1	-1	-1	-1	1
5	-1	-1	-1	1	1
6	0	1	1	-1	0
7	0	-1	0	1	-1
8	-1	-1	-1	-1	-1
9	1	-1	-1	1	-1
10	1	-1	1	1	1
11	-1	-1	1	-1	1

12	1	-1	1	-1	-1
13	-1	1	-1	1	-1
14	1	1	1	1	-1
15	0	-1	-1	0	0
16	-1	0	0	-1	0
17	1	1	1	-1	1
18	1	1	0	0	-1
19	-1	-1	1	0	-1
20	-1	1	1	1	1
21	1	1	-1	1	1
22	1	1	-1	-1	-1
23	0	0	1	0	1
24	0	0	1	1	-1
25	1	0	-1	0	0

Prediction Variance Profile



Prediction Variance Surface



JMP Output

I-Optimal Design

Factors

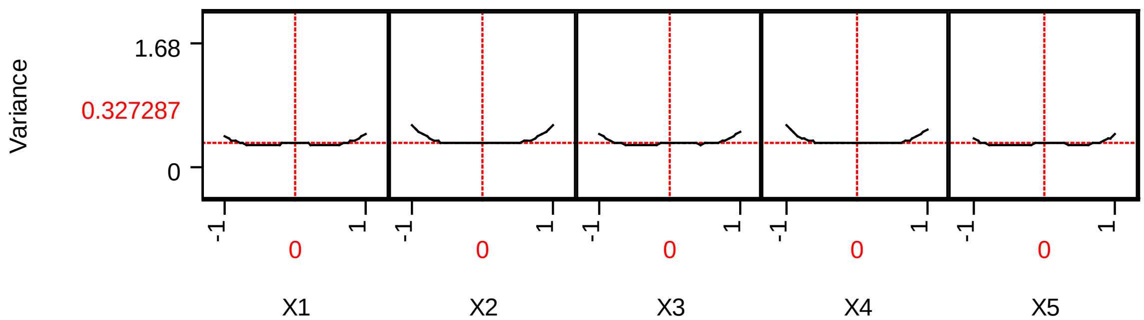
Add N Factors
1

X1 Continuous
X2 Continuous
X3 Continuous
X4 Continuous
X5 Continuous

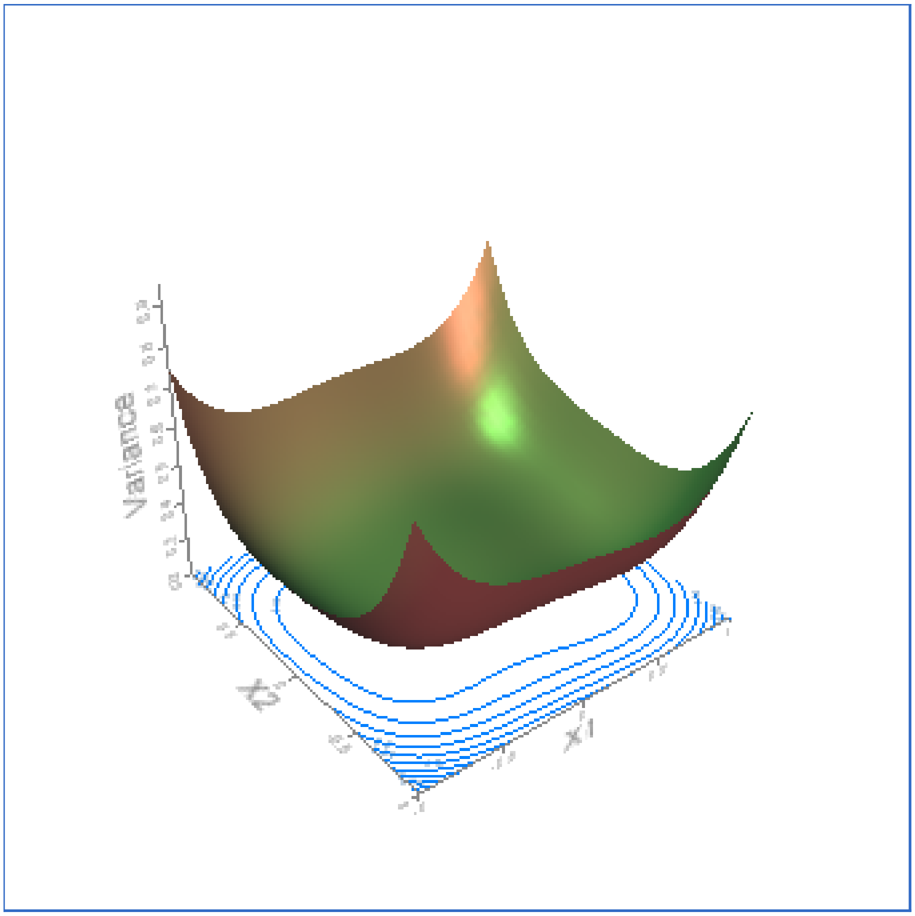
Design

Run	X1	X2	X3	X4	X5
1	-1	0	-1	1	0
2	1	1	1	-1	-1
3	1	1	-1	1	-1
4	0	-1	0	-1	0
5	-1	1	0	0	0
6	0	0	0	1	1
7	-1	0	1	-1	1
8	1	-1	-1	1	1
9	1	0	0	0	-1
10	1	0	-1	-1	0
11	-1	-1	-1	-1	1
12	-1	-1	1	-1	-1
13	1	-1	1	0	1
14	1	1	0	-1	1
15	0	1	1	1	-1
16	-1	0	0	0	-1
17	1	1	1	1	0
18	0	1	-1	0	1
19	1	-1	1	1	-1
20	0	0	0	0	0
21	-1	1	1	1	1
22	-1	1	-1	-1	-1
23	-1	-1	1	1	0
24	0	-1	-1	0	-1
25	0	0	1	0	0

Prediction Variance Profile



Prediction Variance Surface



11.29. Suppose that you want to fit a *second*-order response surface model in a situation where there are k = 4 factors; however, one of the factors is categorical with two levels. What model should you consider for this experiment? Suggest an appropriate design for this situation.

The following model is considered for this experiment:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \gamma z + \delta_1 x_1 z + \delta_2 x_2 z + \delta_3 x_3 z + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{33} x_3^2 + \varepsilon u$$

The following JMP output identifies a computer generated design utilizing the *I*-criterion. Because this is a response surface model which is generally used for prediction, the *I*-criterion is an appropriate choice for a computer generated design.

JMP Output

Custom Design

Factors

Add N Factors
1

X1 Continuous
X2 Continuous
X3 Continuous
X4 Categorical

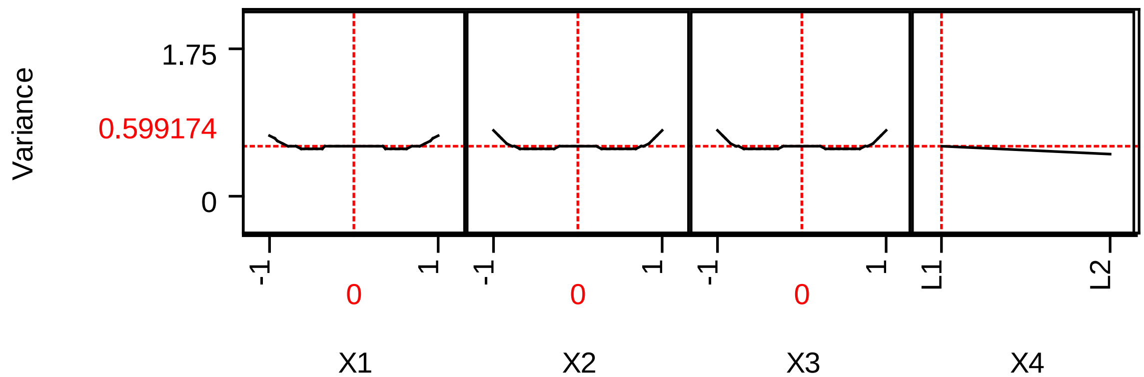
Model

Intercept
Intercept
X1
X2
X3
X4
X1*X2
X1*X3
X1*X4
X2*X3
X2*X4
X3*X4
X1*X1
X2*X2
X3*X3

Design

Run	X1	X2	X3	X4
1	0	0	0	L2
2	-1	-1	-1	L2
3	1	-1	1	L1
4	1	1	1	L2
5	0	-1	-1	L1
6	0	1	1	L1
7	-1	0	1	L1
8	0	1	-1	L2
9	-1	-1	0	L1
10	0	-1	1	L2
11	-1	1	-1	L1
12	1	-1	-1	L2
13	1	0	-1	L1
14	0	0	0	L2
15	1	1	0	L1
16	-1	1	1	L2

Prediction Variance Profile



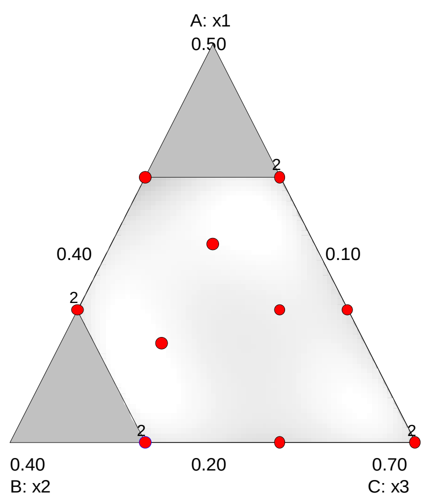
11.30. An experimenter wishes to run a three-component mixture experiment. The constraints in the components proportions are as follows:

$$0.2 \le x_1 \le 0.4$$
$$0.1 \le x_2 \le 0.3$$
$$0.4 \le x_3 \le 0.7$$

(a) Set up an experiment to fit a quadratic mixture model. Use n=14 runs, with 4 replicates. Use the D-criteria.

Std	x1	x2	x3
1	0.2	0.3	0.5
2	0.3	0.3	0.4
3	0.3	0.15	0.55
4	0.2	0.1	0.7
5	0.4	0.2	0.4
6	0.4	0.1	0.5
7	0.2	0.2	0.6
8	0.275	0.25	0.475
9	0.35	0.175	0.475
10	0.3	0.1	0.6
11	0.2	0.3	0.5
12	0.3	0.3	0.4
13	0.2	0.1	0.7
14	0.4	0.1	0.5

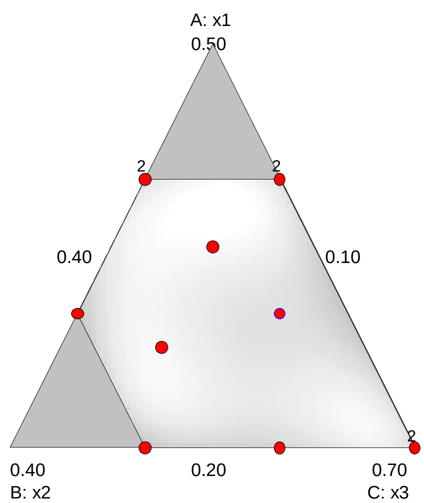
(b) Draw the experimental design region.



(c) Set up an experiment to fit a quadratic mixture model with n=12 runs, assuming that three of these runs are replicated. Use the D-criterion.

Std	x_1	x_2	x_3
1	0.3	0.15	0.55

2	0.2	0.3	0.5
3	0.3	0.3	0.4
4	0.2	0.1	0.7
5	0.4	0.2	0.4
6	0.4	0.1	0.5
7	0.2	0.2	0.6
8	0.275	0.25	0.475
9	0.35	0.175	0.475
10	0.2	0.1	0.7
11	0.4	0.1	0.5
12	0.4	0.2	0.4



(d) Comment on the two designs you have found.

The design points are the same for both designs except that the edge center on the x1-x3 edge is not included in the second design. None of the replicates for either design are in the center of the experimental region. The experimental runs are fairly uniformly spaced in the design region.

11.31. Myers, Montgomery and Anderson-Cook (2009) describe a gasoline blending experiment involving three mixture components. There are no constraints on the mixture proportions, and the following 10 run design is used.

Design Point	x_1	x_2	x_3	y(mpg)
1	1	0	0	24.5, 25.1
2	0	1	0	24.8, 23.9
3	0	0	1	22.7, 23.6
4	½	½	0	25.1
5	½	0	½	24.3
6	0	½	½	23.5
7	1/3	1/3	1/3	24.8, 24.1
8	2/3	1/6	1/6	24.2
9	1/6	2/3	1/6	23.9
10	1/6	1/6	2/3	23.7

(a) What type of design did the experimenters use?

A simplex centroid design was used.

(b) Fit a quadratic mixture model to the data. Is this model adequate?

Design Expert Output

Response: y
ANOVA for Mixture Quadratic Model
Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	4.22	5	0.84	3.90	0.0435	significant
Linear Mixture	*3.92*	*2*	*1.96*	*9.06*	*0.0088*	
AB	*0.15*	*1*	*0.15*	*0.69*	*0.4289*	
AC	*0.081*	*1*	*0.081*	*0.38*	*0.5569*	
BC	*0.077*	*1*	*0.077*	*0.36*	*0.5664*	
Residual	1.73	8	0.22			
Lack of Fit	*0.50*	*4*	*0.12*	*0.40*	*0.8003*	*not significant*
Pure Error	*1.24*	*4*	*0.31*			
Cor Total	5.95	13				

The Model F-value of 3.90 implies the model is significant. There is only a 4.35% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.47	R-Squared	0.7091
Mean	24.16	Adj R-Squared	0.5274
C.V.	1.93	Pred R-Squared	0.1144
PRESS	5.27	Adeq Precision	5.674

Component	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High
A-x1	24.74	1	0.32	24.00	25.49
B-x2	24.31	1	0.32	23.57	25.05
C-x3	23.18	1	0.32	22.43	23.92
AB	1.51	1	1.82	-2.68	5.70
AC	1.11	1	1.82	-3.08	5.30
BC	-1.09	1	1.82	-5.28	3.10

Final Equation in Terms of Pseudo Components:

y =
+24.74 * A
+24.31 * B
+23.18 * C
+1.51 * A * B
+1.11 * A * C
-1.09 * B * C

Final Equation in Terms of Real Components:

y =
+24.74432 * x1
+24.31098 * x2
+23.17765 * x3
+1.51364 * x1 * x2
+1.11364 * x1 * x3
-1.08636 * x2 * x3

The quadratic terms appear to be insignificant. The analysis below is for the linear mixture model:

Design Expert Output

Response: y
ANOVA for Mixture Quadratic Model
Analysis of variance table [Partial sum of squares]

	Sum of		Mean	F

Source	**Squares**	**DF**	**Square**	**Value**	**Prob > F**	
Model	3.92	2	1.96	10.64	0.0027	significant
Linear Mixture	*3.92*	*2*	*1.96*	*10.64*	*0.0027*	
Residual	2.03	11	0.18			
Lack of Fit	*0.79*	*7*	*0.11*	*0.37*	*0.8825*	*not significant*
Pure Error	*1.24*	*4*	*0.31*			
Cor Total	5.95	13				

The Model F-value of 10.64 implies the model is significant. There is only a 0.27% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.43	R-Squared	0.6591
Mean	24.16	Adj R-Squared	0.5972
C.V.	1.78	Pred R-Squared	0.3926
PRESS	3.62	Adeq Precision	8.751

Component	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**
A-x1	24.93	1	0.25	24.38	25.48
B-x2	24.35	1	0.25	23.80	24.90
C-x3	23.19	1	0.25	22.64	23.74

Component	**Effect**	**DF**	**Adjusted Std Error**	**Adjusted Effect=0**	**Approx t for H0 Prob > |t|**
A-x1	1.16	1	0.33	3.49	0.0051
B-x2	0.29	1	0.33	0.87	0.4021
C-x3	-1.45	1	0.33	-4.36	0.0011

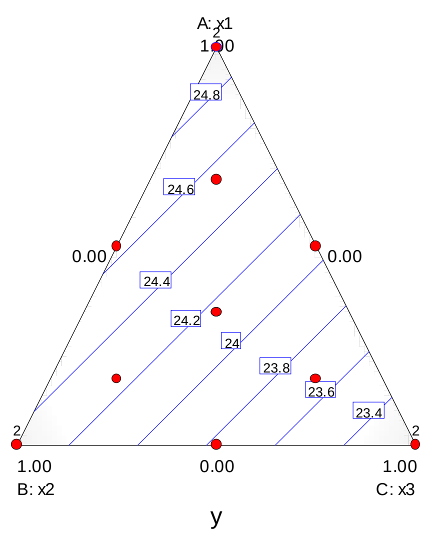
Final Equation in Terms of Pseudo Components:

$$
\begin{aligned}
y &= \\
&+24.93 \quad * A \\
&+24.35 \quad * B \\
&+23.19 \quad * C
\end{aligned}
$$

Final Equation in Terms of Real Components:

$$
\begin{aligned}
y &= \\
&+24.93048 \quad * x1 \\
&+24.35048 \quad * x2 \\
&+23.19048 \quad * x3
\end{aligned}
$$

(c) Plot the response surface contours. What blend would you recommend to maximize the MPG?



To maximize the miles per gallon, the recommended blend is $x_1 = 1$, $x_2 = 0$, and $x_3 = 0$.

11.32. Table P11.9 shows a six-variable RSM design from Jones and Nachtsheim (2011b). Analyze the response data from this experiment.

Table P11.9 – The Design for Problem 11.31

Run(i)	$x_{i.1}$	$x_{i.2}$	$x_{i.3}$	$x_{i.4}$	$x_{i.5}$	$x_{i.6}$	y_i
1	0	1	-1	-1	-1	-1	21.04
2	0	-1	1	1	1	1	10.48
3	1	0	-1	1	1	-1	17.89
4	-1	0	1	-1	-1	1	10.07
5	-1	-1	0	1	-1	-1	7.74
6	1	1	0	-1	1	1	21.01
7	-1	1	1	0	1	-1	16.53
8	1	-1	-1	0	-1	1	20.38
9	1	-1	1	-1	0	-1	8.62
10	-1	1	-1	1	0	1	7.8
11	1	1	1	1	-1	0	23.56
12	-1	-1	-1	-1	1	0	15.24
13	0	0	0	0	0	0	19.91

Design Expert Output

Response 1 Yi

ANOVA for Response Surface Reduced 2FI Model

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

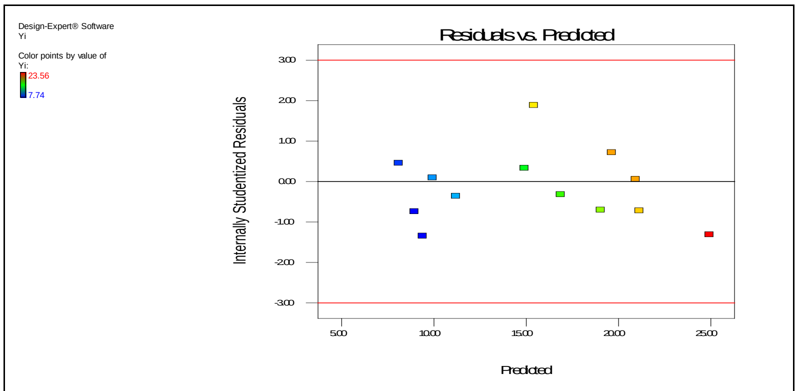
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	366.59	7	52.37	8.46	0.0159	significant
A-Xi.1	*116.14*	*1*	*116.14*	*18.77*	*0.0075*	
B-Xi.2	*75.52*	*1*	*75.52*	*12.20*	*0.0174*	
C-Xi.3	*17.13*	*1*	*17.13*	*2.77*	*0.1570*	
E-Xi.5	*0.27*	*1*	*0.27*	*0.043*	*0.8431*	
AB	*84.75*	*1*	*84.75*	*13.70*	*0.0140*	
AC	*64.42*	*1*	*64.42*	*10.41*	*0.0233*	
AE	*109.00*	*1*	*109.00*	*17.61*	*0.0085*	
Residual	30.94	5	6.19			
Cor Total	397.53	12				

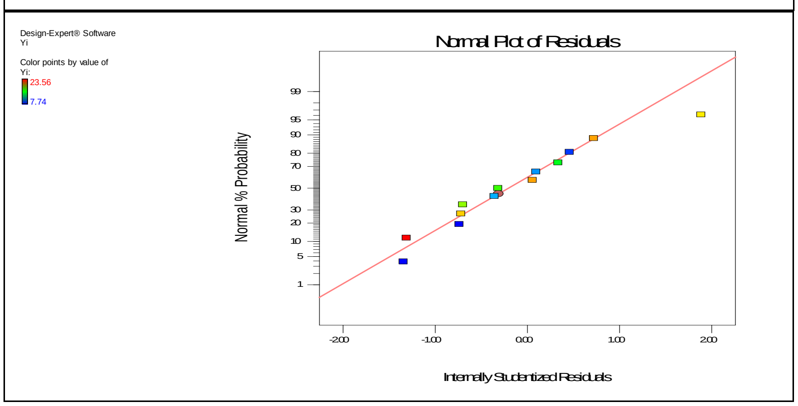
Std. Dev.	2.49	R-Squared	0.9222
Mean	15.41	Adj R-Squared	0.8132
C.V. %	16.15	Pred R-Squared	0.5242
PRESS	189.14	Adeq Precision	8.627

Factor	**Coefficient Estimate**	**df**	**Error**	**Standard Low**	**95% CI High**	**95% CI VIF**
Intercept	15.41	1	0.69	13.63	17.18	
A-Xi.1	3.41	1	0.79	1.39	5.43	1.00
B-Xi.2	2.75	1	0.79	0.73	4.77	1.00
C-Xi.3	-1.31	1	0.79	-3.33	0.71	1.00
E-Xi.5	-0.16	1	0.79	-2.19	1.86	1.00
AB	3.57	1	0.96	1.09	6.04	1.20
AC	-3.11	1	0.96	-5.59	-0.63	1.20
AE	-4.04	1	0.96	-6.52	-1.57	1.20

Final Equation in Terms of Actual Factors:

Yi	=
+15.40538	
+3.40800	* Xi.1
+2.74800	* Xi.2
-1.30900	* Xi.3
-0.16400	* Xi.5
+3.56550	* Xi.1 * Xi.2
-3.10850	* Xi.1 * Xi.3
-4.04350	* Xi.1 * Xi.5





The main effects A and B are significant as well as the AB, AC and AE two-factor interactions. Factors C and E are kept in the model to maintain hierarchy. The normal probability plot of residuals shows a possible outlier in the upper right corner.

11.33. An article in *Quality Progress* ("For Starbucks, It's in the Bag," March 2011, pp. 18-23) describes using a central composite design to improve the packaging of one-pound coffee. The objective is to produce an airtight seal that is easy to open without damaging the top of the coffee bag. The experimenters studied the three factors – x_1 = plastic viscosity (300-400 centipoise), x_2 = clamp pressure (170-190 psi), and x_3 = plate gap (-3, +3 mm) and two responses – y_1 = tear and y_2 = leakage. The design is shown in Table P11.10. The tear response was measured on a scale from 0 to 9 (good to bad) and the leakage was proportion failing. Each run used a sample of 20 bags for response measurement.

Table P11.10 – The Coffee Bag Experiment in Problem 11.32

Run	Viscosity	Pressure	Plate gap	Tear	Leakage
Center	350	180	0	0	0.15
Axial	350	170	0	0	0.5

Factorial	319	186	1.8	0.45	0.15
Factorial	380	174	1.8	0.85	0.05
Center	350	180	0	0.35	0.15
Axial	300	180	0	0.3	0.45
Axial	400	180	0	0.7	0.25
Axial	350	190	0	1.9	0
Center	350	180	0	0.25	0.05
Factorial	319	186	-1.8	0.1	0.35
Factorial	380	186	-1.8	0.15	0.4
Axial	350	180	3	3.9	0
Factorial	380	174	-1.8	0	0.45
Center	350	180	0	0.55	0.2
Axial	350	180	-3	0	1
Factorial	319	174	-1.8	0.05	0.2
Factorial	319	174	1.8	0.4	0.25
Factorial	380	186	1.8	4.3	0.05
Center	350	180	0	0	0

(a) Build a second-order model for the tear response.

Design Expert Output

Response 1 **Tear**
ANOVA for Response Surface Reduced Cubic 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	26.40	8	3.30	13.72	0.0002	significant
A-Viscosity	*1.81*	*1*	*1.81*	*7.52*	*0.0207*	
B-Pressure	*3.48*	*1*	*3.48*	*14.47*	*0.0035*	
C-Plate Gap	*11.00*	*1*	*11.00*	*45.73*	*< 0.0001*	
AB	*1.53*	*1*	*1.53*	*6.36*	*0.0303*	
AC	*2.31*	*1*	*2.31*	*9.60*	*0.0113*	
BC	*1.36*	*1*	*1.36*	*5.66*	*0.0387*	
C2	*3.54*	*1*	*3.54*	*14.72*	*0.0033*	
ABC	*1.36*	*1*	*1.36*	*5.66*	*0.0387*	
Residual	2.41	10	0.24			
Lack of Fit	*2.18*	*6*	*0.36*	*6.53*	*0.0454*	*significant*
Pure Error	*0.22*	*4*	*0.056*			
Cor Total	28.81	18				

Std. Dev.	0.49	R-Squared	0.9165
Mean	0.75	Adj R-Squared	0.8497
C.V. %	65.41	Pred R-Squared	0.4004
PRESS	17.27	Adeq Precision	14.560

Factor	**Coefficient Estimate**	**df**	**Error**	**Standard Low**	**95% CI High**	**95% CI VIF**
Intercept	0.39	1	0.15	0.065	0.72	
A-Viscosity	0.36	1	0.13	0.068	0.66	1.00
B-Pressure	0.50	1	0.13	0.21	0.80	1.00
C-Plate Gap	0.90	1	0.13	0.60	1.19	1.00
AB	0.44	1	0.17	0.051	0.82	1.00
AC	0.54	1	0.17	0.15	0.92	1.00
BC	0.41	1	0.17	0.026	0.80	1.00
C2	0.50	1	0.13	0.21	0.79	1.00
ABC	0.41	1	0.17	0.026	0.80	1.00

Final Equation in Terms of Actual Factors:

Tear =	
+131.47081	
-0.41839	* Viscosity
-0.75140	* Pressure
+68.98264	* Plate Gap
+2.39071E-003	* Viscosity * Pressure
-0.21562	* Viscosity * Plate Gap
-0.39948	* Pressure * Plate Gap
+0.15424	* Plate Gap^2
+1.25228E-003	* Viscosity * Pressure * Plate Gap

The second-order model for the tear response includes all three main effects, all two-factor interactions, the three-factor interaction and B^2.

(b) Build a second-order model for the leakage response.

Design Expert Output

Response 2 Leakage
ANOVA for Response Surface Reduced Quadratic 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	0.62	2	0.31	10.94	0.0010	significant
C-Plate Gap	*0.49*	*1*	*0.49*	*17.13*	*0.0008*	
C2	*0.14*	*1*	*0.14*	*4.75*	*0.0445*	
Residual	0.46	16	0.028			
Lack of Fit	*0.43*	*12*	*0.036*	*5.30*	*0.0603*	*not significant*
Pure Error	*0.027*	*4*	*6.750E-003*			
Cor Total	1.08	18				

Std. Dev.	0.17	R-Squared	0.5776
Mean	0.24	Adj R-Squared	0.5248
C.V. %	68.98	Pred R-Squared	0.1949
PRESS	0.87	Adeq Precision	10.222

Factor	**Coefficient Estimate**	**df**	**Error**	**Standard Low**	**95% CI High**	**95% CI VIF**
Intercept	0.17	1	0.050	0.068	0.28	
C-Plate Gap	-0.19	1	0.046	-0.29	-0.092	1.00
C2	0.098	1	0.045	2.704E-003	0.19	1.00

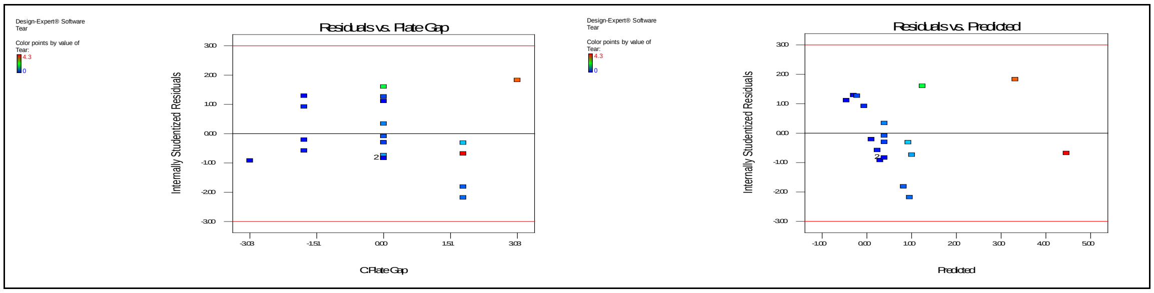
Final Equation in Terms of Actual Factors:

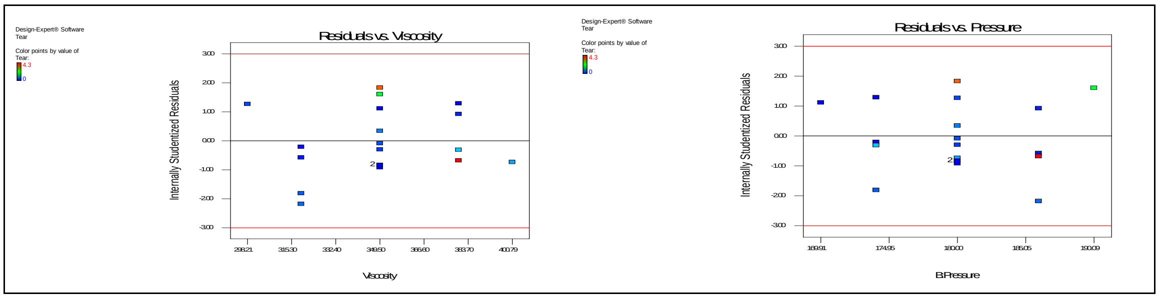
Leakage =	
+0.17449	
-0.10503	* Plate Gap
+0.030162	* Plate Gap2

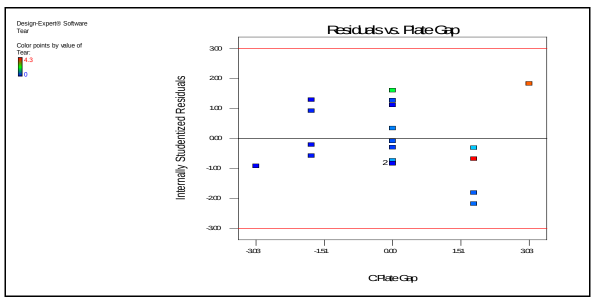
For the Leakage response, only variable C – Plate Gap and C^2 are important.

(c) Analyze the residuals for both models. Do transformations seem necessary for either response? If so, refit the models in the transformed metric.

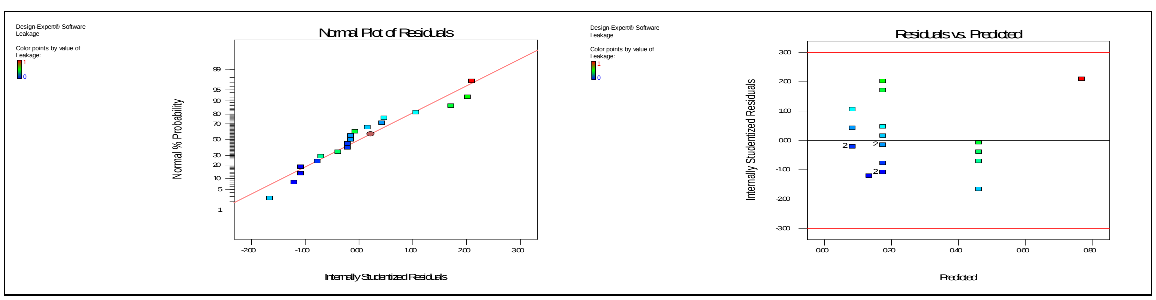
For the Tear response there is a little bit of non-linearity in the C response.

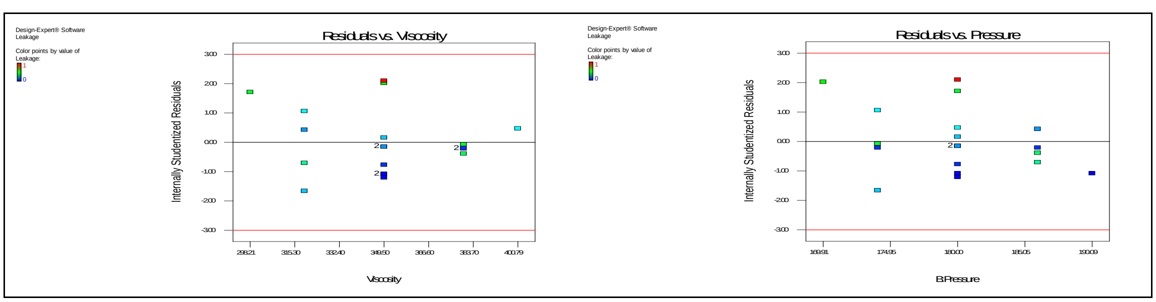


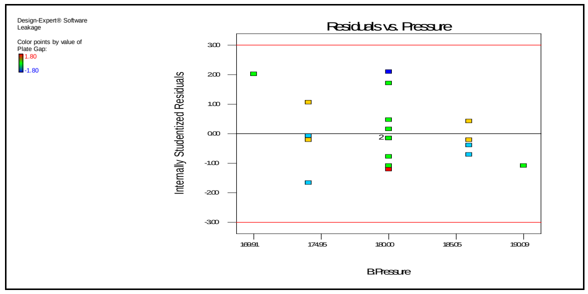




For the Tear response the residuals look fine. No concerns with model adequacy.



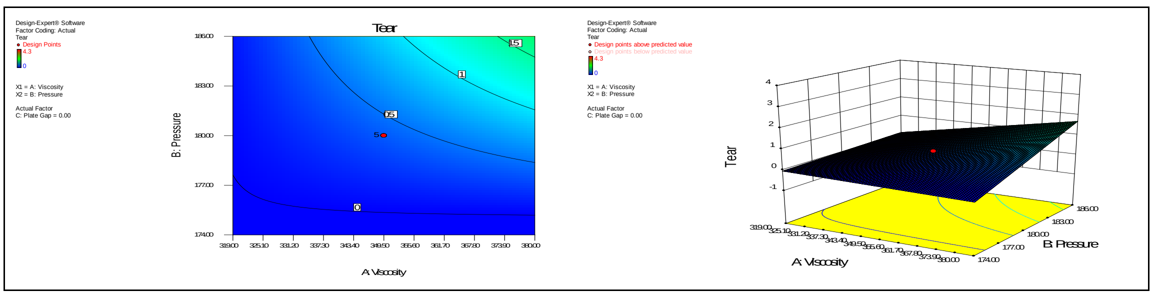




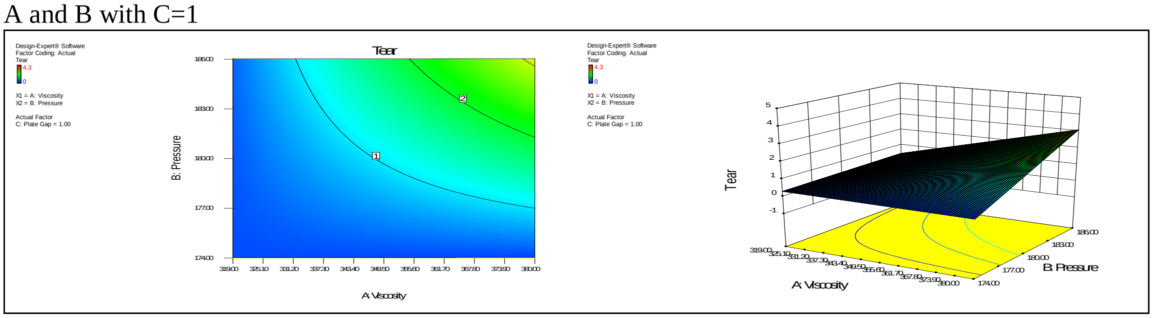
(d) Construct response surface plots and contour plots for both responses. Provide interpretations for the fitted surfaces.

For the Tear Response:

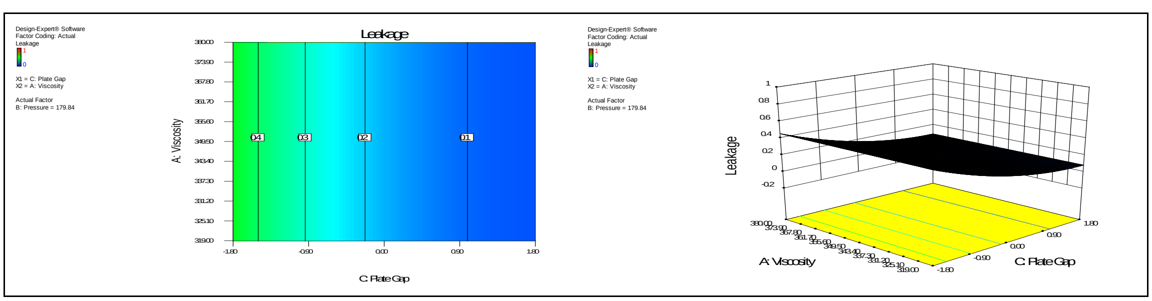
A and B with C=0



A and B with C=1

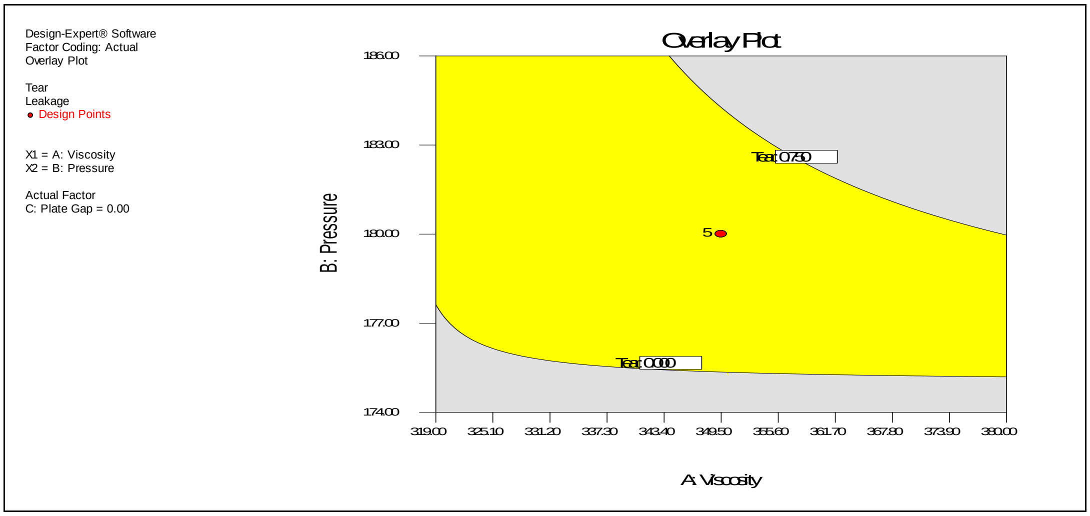


For the Leakage Response, factors A and C, with B=0



(e) What conditions would you recommend for process optimization to minimize leakage and keep tear below 0.75.

Recommend running the process with C<0. When C gets >0 then there is very little range where the Tear response is less than 0.75. The graph below shows the A and B ranges when C=0.



11.34. Box and Liu (1999) describe an experiment flying paper helicopters where the objective is to maximize flight time. They used a central composite design shown in Table P11.11. Each run involved a single helicopter made to the following specifications: x_1 = wing area (in^2), -1 = 11.80 and +1 = 13.00; x_2 = wing-length to width ratio, -1 = 2.25 and +1 = 2.78; x_3 = base width (in), -1 = 1.00 and +1 = 1.50; and x_4 = base length (in), -1 = 1.50 and +1 = 2.50. Each helicopter was flown four times and the average flight time and the standard deviation of flight time was recorded.

Table P11.11- The Paper Helicopter Experiment

Std Ord	Run Ord	Wing Area	Wing Ratio	Base Width	Base Length	Avg Flight Time	St Dev Flight Time
1	9	-1	-1	-1	-1	3.67	0.052
2	21	1	-1	-1	-1	3.69	0.052
3	14	-1	1	-1	-1	3.74	0.055
4	4	1	1	-1	-1	3.7	0.062
5	2	-1	-1	1	-1	3.72	0.052
6	19	1	-1	1	-1	3.55	0.065
7	22	-1	1	1	-1	3.97	0.052
8	25	1	1	1	-1	3.77	0.098
9	27	-1	-1	-1	1	3.5	0.079
10	13	1	-1	-1	1	3.73	0.072
11	20	-1	1	-1	1	3.58	0.083
12	6	1	1	-1	1	3.63	0.132
13	12	-1	-1	1	1	3.44	0.058
14	17	1	-1	1	1	3.55	0.049
15	26	-1	1	1	1	3.7	0.081
16	1	1	1	1	1	3.62	0.051
17	8	-2	0	0	0	3.61	0.129
18	15	2	0	0	0	3.64	0.085
19	7	0	-2	0	0	3.55	0.1
20	5	0	2	0	0	3.73	0.063

21	29	0	0	-2	0	3.61	0.051
22	28	0	0	2	0	3.6	0.095
23	16	0	0	0	-2	3.8	0.049
24	18	0	0	0	2	3.6	0.055
25	24	0	0	0	0	3.77	0.032
26	10	0	0	0	0	3.75	0.055
27	23	0	0	0	0	3.7	0.072
28	11	0	0	0	0	3.68	0.055
29	3	0	0	0	0	3.69	0.078
30	30	0	0	0	0	3.66	0.058

(a) Fit a second-order model to the average flight time response. The flight time response is driven by B and D, several two-factor interactions and A^2 and C^2 terms. A and C are in the model to maintain hierarchy.

Design Expert Output

Response 1 Avg Flt Time
ANOVA for Response Surface Reduced Quadratic 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	0.28	11	0.025	15.78	< 0.0001	significant
A-Wing Area	*1.667E-005*	*1*	*1.667E-005*	*0.010*	*0.9198*	
B-Length/Width	*0.062*	*1*	*0.062*	*38.83*	*< 0.0001*	
C-Base Width	*1.500E-004*	*1*	*1.500E-004*	*0.094*	*0.7628*	
D-Base Length	*0.089*	*1*	*0.089*	*55.61*	*< 0.0001*	
AB	*0.013*	*1*	*0.013*	*8.28*	*0.0100*	
AC	*0.023*	*1*	*0.023*	*14.09*	*0.0015*	
AD	*0.031*	*1*	*0.031*	*19.17*	*0.0004*	
BC	*0.034*	*1*	*0.034*	*21.43*	*0.0002*	
CD	*7.225E-003*	*1*	*7.225E-003*	*4.52*	*0.0475*	
A2	*7.511E-003*	*1*	*7.511E-003*	*4.70*	*0.0438*	
C2	*0.013*	*1*	*0.013*	*8.04*	*0.0110*	
Residual	*0.029*	*18*	*1.597E-003*			
Lack of Fit	*0.020*	*13*	*1.513E-003*	*0.83*	*0.6382*	*not significant*
Pure Error	*9.083E-003*	*5*	*1.817E-003*			
Cor Total	*0.31*	*29*				

Std. Dev.	0.040	R-Squared	0.9060
Mean	3.66	Adj R-Squared	0.8486
C.V. %	1.09	Pred R-Squared	0.7581
PRESS	0.074	Adeq Precision	19.913

Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	3.70	1	0.012	3.67	3.72	
A-Wing Area	-8.333E-004	1	8.158E-003	-0.018	0.016	1.00
B-Length/Width	0.051	1	8.158E-003	0.034	0.068	1.00
C-Base Width	2.500E-003	1	8.158E-003	-0.015	0.020	1.00
D-Base Length	-0.061	1	8.158E-003	-0.078	-0.044	1.00
AB	-0.029	1	9.991E-003	-0.050	-7.759E-003	1.00
AC	-0.037	1	9.991E-003	-0.058	-0.017	1.00
AD	0.044	1	9.991E-003	0.023	0.065	1.00
BC	0.046	1	9.991E-003	0.025	0.067	1.00
CD	-0.021	1	9.991E-003	-0.042	-2.590E-004	1.00
A2	-0.016	1	7.493E-003	-0.032	-5.068E-004	1.01
C2	-0.021	1	7.493E-003	-0.037	-5.507E-003	1.01

Final Equation in Terms of Actual Factors:

Avg Flt Time	**=**
+3.70	
-8.333E-004	* A
+0.051	* B
+2.500E-003	* C
-0.061	* D
-0.029	* A * B
-0.037	* A * C
+0.044	* A * D
+0.046	* B * C
-0.021	* C * D
-0.016	* A2
-0.021	* C2

Final Equation in Terms of Actual Factors:

Avg Flt Time	**=**
-8.13881	
+1.59365	* Wing Area
+1.56132	* Length/Width Ratio
+2.54425	* Base Width
-1.71750	* Base Length
-0.18082	* Wing Area * Length/Width Ratio
-0.25000	* Wing Area * Base Width
+0.14583	* Wing Area * Base Length
+0.69811	* Length/Width Ratio * Base Width
-0.17000	* Base Width * Base Length
-0.045139	* Wing Area2
-0.34000	* Base Width2

(b) Fit a second-order model to the standard deviation of flight time response. The standard deviation response is driven by the CD interaction and A^2. A, C and D are in the model to maintain hierarchy.

Design Expert Output

Response 2 Stdev Flt Time
ANOVA for Response Surface Reduced Quadratic 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	5.562E-003	5	1.112E-003	2.57	0.0533	not significant
A-Wing Area	*1.504E-005*	*1*	*1.504E-005*	*0.035*	*0.8536*	
C-Base Width	*2.042E-006*	*1*	*2.042E-006*	*4.72E-003*	*0.9458*	
D-Base Length	*6.934E-004*	*1*	*6.934E-004*	*1.60*	*0.2176*	
CD	*1.871E-003*	*1*	*1.871E-003*	*4.32*	*0.0484*	
A2	*2.981E-003*	*1*	*2.981E-003*	*6.89*	*0.0148*	
Residual	0.010	24	4.325E-004			
Lack of Fit	*9.091E-003*	*19*	*4.785E-004*	*1.86*	*0.2556*	*not significant*
Pure Error	*1.289E-003*	*5*	*2.579E-004*			
Cor Total	0.016	29				

Std. Dev.	0.021	R-Squared	0.3489
Mean	0.069	Adj R-Squared	0.2132
C.V. %	30.14	Pred R-Squared	-0.1717
PRESS	0.019	Adeq Precision	5.702

Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.061	1	4.90E-003	0.051	0.071	
A-Wing Area	-7.917E-004	1	4.24E-003	-9.55E-003	7.970E-003	1.00
C-Base Width	2.917E-004	1	4.24E-003	-8.47E-003	9.053E-003	1.00

D-Base Length	5.375E-003	1	4.24E-003 -3.38E-003	0.014	1.00
CD	-0.011	1	5.19E-003 -0.022	-8.191E-005	1.00
A2	0.010	1	3.87E-003 2.175E-003	0.018	1.00

Final Equation in Terms of Coded Factors:

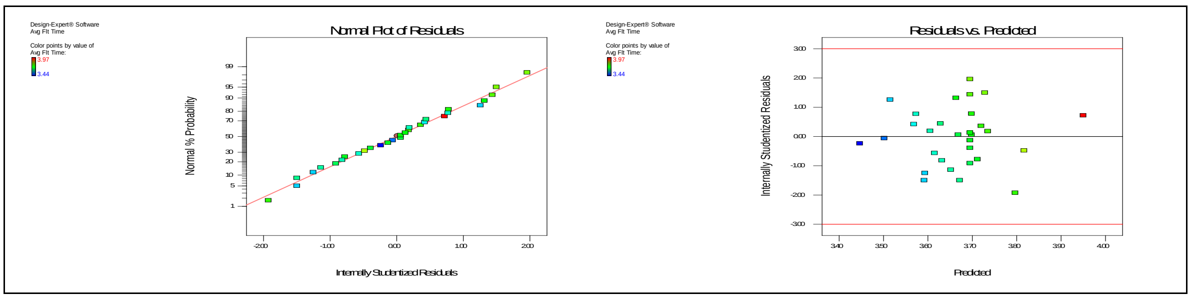
Stdev Flt Time	=
+0.061	
-7.917E-004	* A
+2.917E-004	* C
+5.375E-003	* D
-0.011	* C * D
+0.010	* A2

Final Equation in Terms of Actual Factors:

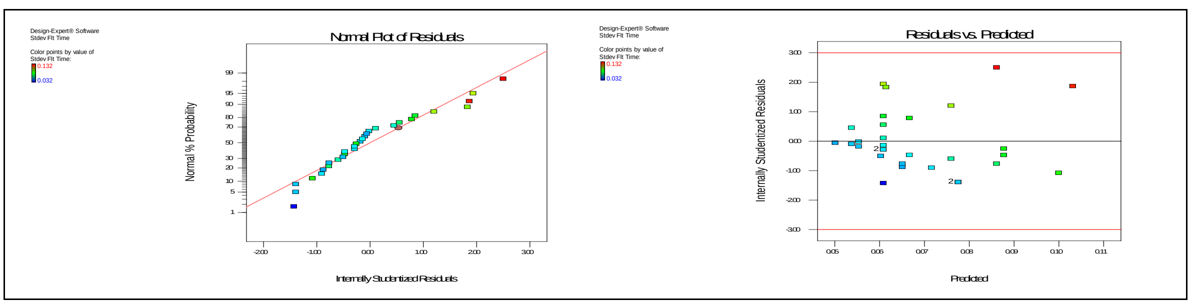
Stdev Flt Time	=
+4.18328	
-0.70217	* Wing Area
+0.17417	* Base Width
+0.11887	* Base Length
-0.086500	* Base Width * Base Length
+0.028260	* Wing Area2

(c) Analyze the residuals for both models from parts (a) and (b). Are transformations on the response(s) necessary? If so, fit the appropriate models.

For the average flight time response the residuals look fine. The residual by factor are also okay.



For the standard deviation of flight time response the plot of residuals vs. predicted shows an inequality of variance. Recommend doing a Log transformation.



Design Expert Output

Response 2 **Stdev Flt Time**
Transform: **Base 10 Log** **Constant: 0**
ANOVA for Response Surface Reduced Quadratic 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	0.18	5	0.035	2.34	0.0725	not sig

A-Wing Area	*1.089E-004*	*1*	*1.089E-004*	*7.212E-003*	*0.9330*	
C-Base Width	*3.136E-004*	*1*	*3.136E-004*	*0.021*	*0.8866*	
D-Base Length	*0.024*	*1*	*0.024*	*1.61*	*0.2167*	
CD	*0.062*	*1*	*0.062*	*4.11*	*0.0539*	
A2	*0.090*	*1*	*0.090*	*5.96*	*0.0224*	
Residual	0.36	24	0.015			
Lack of Fit	*0.27*	*19*	*0.014*	*0.77*	*0.6920*	*not sig*
Pure Error	*0.092*	*5*	*0.018*			
Cor Total	0.54	29				

Std. Dev.	0.12	R-Squared	0.3278
Mean	-1.18	Adj R-Squared	0.1878
C.V. %	10.39	Pred R-Squared	-0.0669
PRESS	0.58	Adeq Precision	5.304

Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	-1.23	1	0.029	-1.29	-1.17	
A-Wing Area	-2.130E-003	1	0.025	-0.054	0.050	1.00
C-Base Width	3.615E-003	1	0.025	-0.048	0.055	1.00
D-Base Length	0.032	1	0.025	-0.020	0.084	1.00
CD	-0.062	1	0.031	-0.13	1.13E-003	1.00
A2	0.056	1	0.023	8.64E-003	0.10	1.00

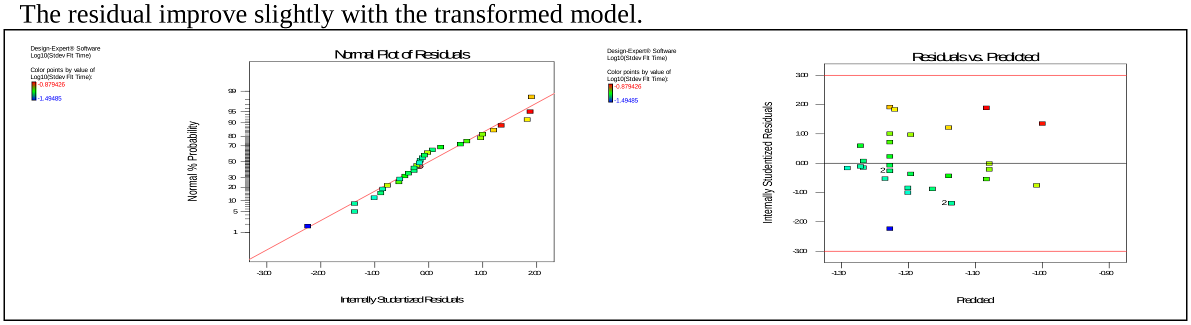
Final Equation in Terms of Coded Factors:

Log10(Stdev Flt Time) =
-1.23
-2.130E-003 * A
+3.615E-003 * C
+0.032 * D
-0.062 * C * D
+0.056 * A2

Final Equation in Terms of Actual Factors:

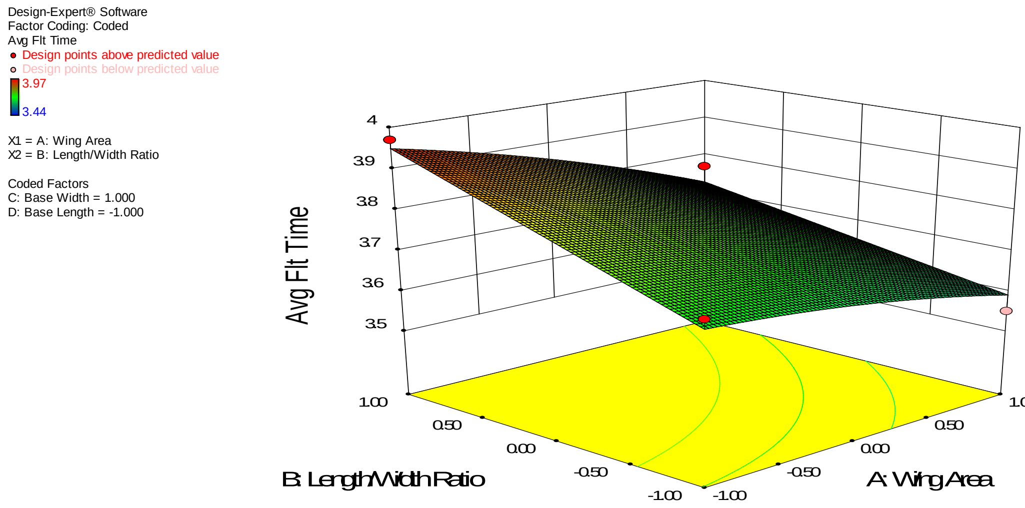
Log10(Stdev Flt Time) =
+21.30149
-3.85446 * Wing Area
+1.01078 * Base Width
+0.68634 * Base Length
-0.49816 * Base Width * Base Length
+0.15528 * Wing Area2

The residual improve slightly with the transformed model.

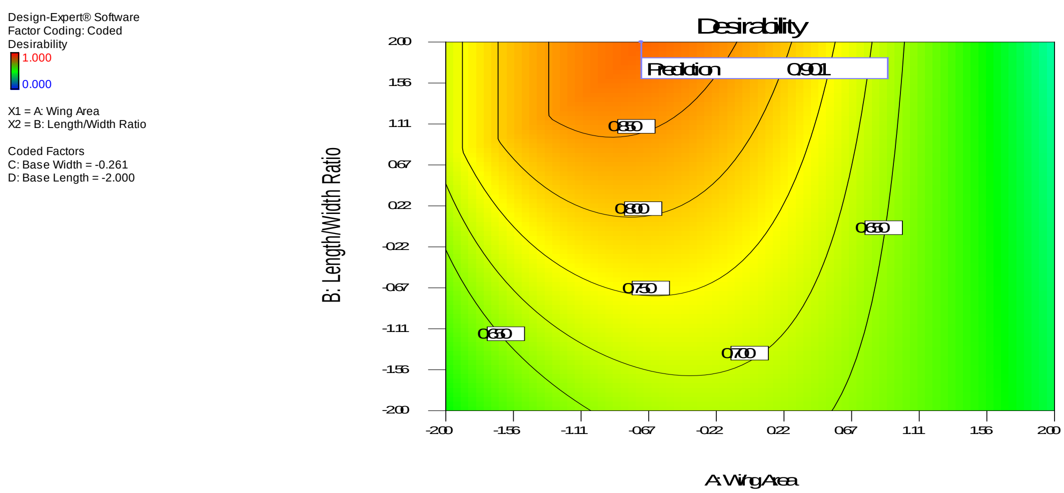


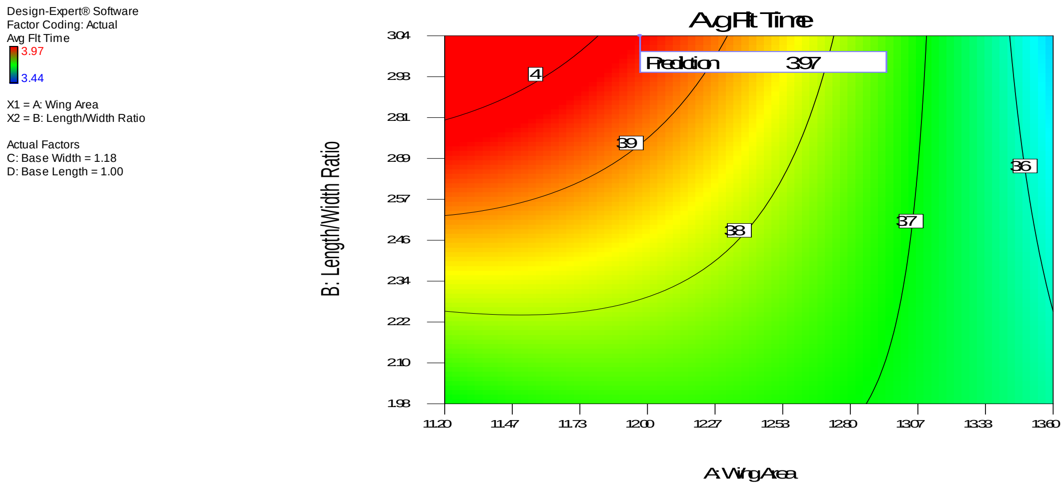
(d) What design would you recommend to maximize the flight time?

Run A at the low level, B at the high level, C at the high level and D at the low level. This give a flight time of almost 4 seconds.



(e) What design would you recommend to maximize the flight time while simultaneously minimizing the standard deviation of flight time? Running A at -0.715, B at -2, C at -0.261 and D at -2, in coded terms. Applying these values to the regression equations estimates the hang time of 3.97 seconds and the standard deviation at 0.051.

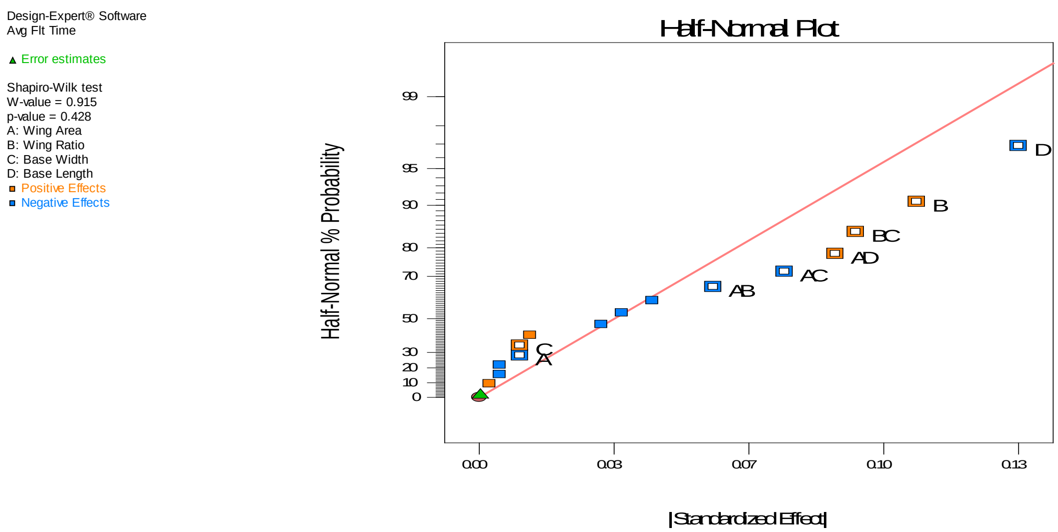




11.35. The Paper Helicopter Experiment Revisited. Reconsider the paper helicopter experiment in Problem 11.34. This experiment was actually run in two blocks. Block 1 consisted of the first 16 runs in Table P11.11 (standard order runs 1-16) and two center points (standard order runs 25 and 26).

(a) Fit the main-effects plus two-factor interaction models to the block 1 data, using both responses.

For the average flight time: The main effects of B and D are significant as well as the two-factor interactions of AB, AC, AD, BC. A and C are added to maintain hieracrcy.



Design Expert Output

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	**0.22**	**8**	**0.027**	**12.99**	**0.0008**	**significant**
A-Wing Area	*4.00E-004*	*1*	*4.00E-004*	*0.19*	*0.6738*	
B-Wing Ratio	*0.046*	*1*	*0.046*	*22.04*	*0.0016*	
C-Base Width	*4.00E-004*	*1*	*4.00E-004*	*0.19*	*0.6738*	
D-Base Length	*0.070*	*1*	*0.070*	*33.49*	*0.0004*	
AB	*0.013*	*1*	*0.013*	*6.31*	*0.0363*	

AC	*0.022*	*1*	*0.022*	*10.73*	*0.0113*	
AD	*0.031*	*1*	*0.031*	*14.61*	*0.0051*	
BC	*0.034*	*1*	*0.034*	*16.32*	*0.0037*	
Curvature	0.018	1	0.018	8.48	0.0195	significant
Residual	0.017	8	2.09E-003			
Lack of Fit	*0.017*	*7*	*2.36E-003*	*11.84*	*0.2202*	*not significant*
Pure Error	*2.00E-004*	*1*	*2.00E-004*			
Cor Total	0.25	17				

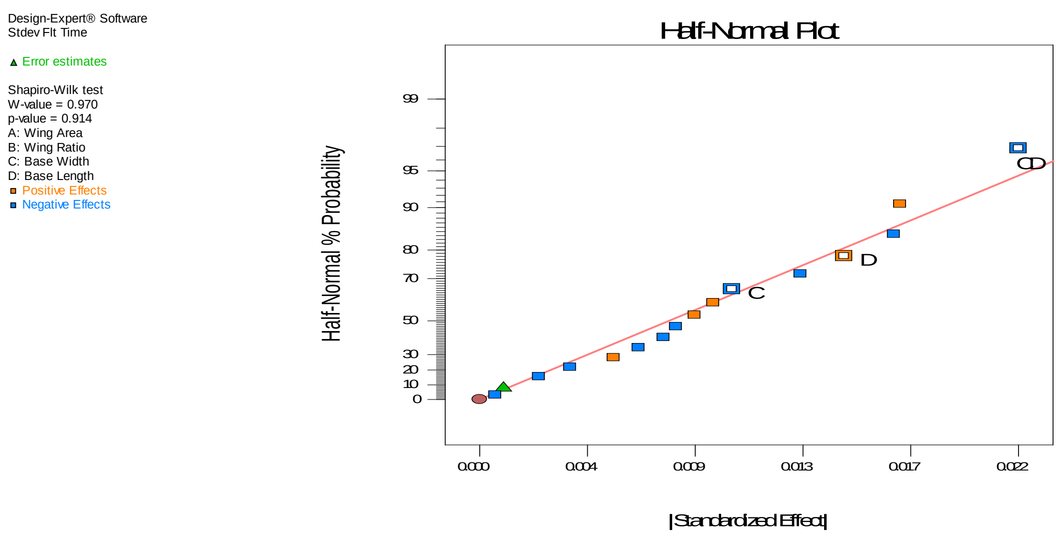
Std. Dev.	0.062	R-Squared	0.8631
Mean	3.67	Adj R-Squared	0.7414
C.V. %	1.69	Pred R-Squared	0.5568
PRESS	0.11	Adeq Precision	10.899

Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	3.67	1	0.015	3.64	3.70	
A-Wing Area	-5.00E-003	1	0.015	-0.040	0.030	1.00
B-Wing Ratio	0.054	1	0.015	0.019	0.089	1.00
C-Base Width	5.00E-003	1	0.015	-0.030	0.040	1.00
D-Base Length	-0.066	1	0.015	-0.10	-0.031	1.00
AB	-0.029	1	0.015	-0.064	6.29E-003	1.00
AC	-0.037	1	0.015	-0.073	-2.46E-003	1.00
AD	0.044	1	0.015	8.71E-003	0.079	1.00
BC	0.046	1	0.015	0.011	0.081	1.00

Final Equation in Terms of Coded Factors:

Avg Flt Time	=
+3.67	
-5.00E-003	* A
+0.054	* B
+5.00E-003	* C
-0.066	* D
-0.029	* A * B
-0.037	* A * C
+0.044	* A * D
+0.046	* B * C

For the Standard Deviation Flight Time response: Only the CD interaction is significant. C and D are added to maintain hierarchy.



Design Expert Output

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	3.136E-003	3	1.045E-003	2.93	0.0735	not significant
C-Base Width	*4.101E-004*	*1*	*4.101E-004*	*1.15*	*0.3034*	
D-Base Length	*8.556E-004*	*1*	*8.556E-004*	*2.40*	*0.1456*	
CD	*1.871E-003*	*1*	*1.871E-003*	*5.24*	*0.0395*	
Curvature	1.095E-003	1	1.095E-003	3.07	0.1035	not significant
Residual	4.642E-003	13	3.571E-004			
Lack of Fit	*4.377E-003*	*12*	*3.648E-004*	*1.38*	*0.5889*	*not significant*
Pure Error	*2.645E-004*	*1*	*2.645E-004*			
Cor Total	8.872E-003	17				

Std. Dev.	0.020	R-Squared	0.3535
Mean	0.066	Adj R-Squared	0.2149
C.V. %	30.88	Pred R-Squared	-0.0413
PRESS	9.239E-003	Adeq Precision	3.799

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	0.066	1	4.771E-003	0.055	0.076	
C-Base Width	-5.063E-003	1	5.060E-003	-0.016	5.791E-003	1.00
D-Base Length	7.312E-003	1	5.060E-003	-3.54E-003	0.018	1.00
CD	-0.011	1	5.060E-003	-0.022	4.111E-005	1.00

Final Equation in Terms of Coded Factors:

Stdev Flt Time	**=**
+0.066	
-5.063E-003	* C
+7.312E-003	* D
-0.011	* C * D

(b) For the models in part (a) use the two center points to test for lack of fit. Is there an indication that second-order terms are needed? The ANOVA for Average Flight Time shows curvature to be significant. The ANOVA for Standard Deviation Flight Time does not show curvature.

(c) Now use the data from block 2 (standard order runs 17-24 and the remaining center points, standard order runs 27-30) to augment block 1 and fit second-order models to both responses. Check the adequacy of the fit for both models. Does blocking seem to have been important in this experiment? Blocking has little effect on the models.

For the average flight time of the combined design: The main effects of B and D are significant as well as the two-factor interactions of AB, AC, AD, BC and CD. The quadratic terms of A^2, B^2 and C^2 are significant. A and C are added to maintain hierarchy.

Design Expert Output

Response 1 Avg Flt Time
ANOVA for Response Surface Reduced Quadratic Model
Analysis of variance table [Partial sum of squares - Type III]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F	
Block	1.68E-003	1	1.681E-003			
Model	0.29	12	0.024	22.14	< 0.0001	significant
A-Wing Area	*1.67E-005*	*1*	*1.667E-005*	*0.015*	*0.9027*	
B-Wing Ratio	*0.062*	*1*	*0.062*	*57.40*	*< 0.0001*	
C-Base Width	*1.50E-004*	*1*	*1.500E-004*	*0.14*	*0.7143*	
D-Base Length	*0.089*	*1*	*0.089*	*82.21*	*< 0.0001*	
AB	*0.013*	*1*	*0.013*	*12.24*	*0.0030*	
AC	*0.022*	*1*	*0.022*	*20.83*	*0.0003*	
AD	*0.031*	*1*	*0.031*	*28.35*	*< 0.0001*	
BC	*0.034*	*1*	*0.034*	*31.68*	*< 0.0001*	

CD	*7.22E-003*	*1*	*7.225E-003*	*6.69*	*0.0199*	
A2	*0.011*	*1*	*0.011*	*10.26*	*0.0055*	
B2	*7.33E-003*	*1*	*7.336E-003*	*6.79*	*0.0191*	
C2	*0.017*	*1*	*0.017*	*16.00*	*0.0010*	
Residual	0.017	16	1.080E-003			
Lack of Fit	*0.016*	*12*	*1.351E-003*	*5.03*	*0.0659*	*not significant*
Pure Error	*1.07E-003*	*4*	*2.688E-004*			
Cor Total	0.31	29				

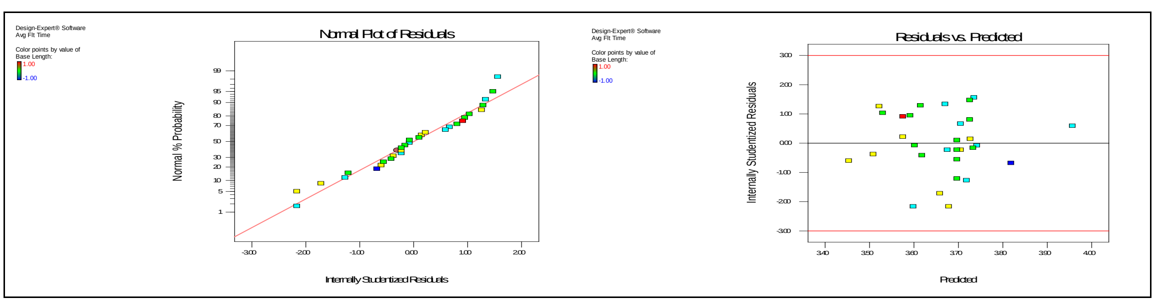
Std. Dev.	0.033	R-Squared	0.9432
Mean	3.66	Adj R-Squared	0.9006
C.V. %	0.90	Pred R-Squared	0.7728
PRESS	0.069	Adeq Precision	22.417

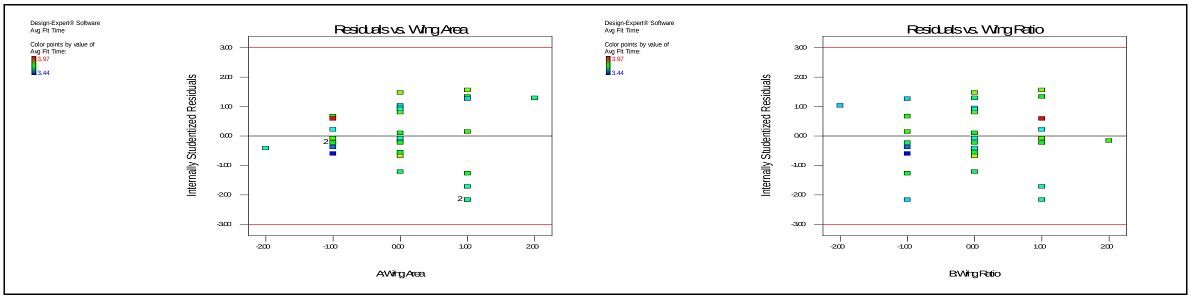
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	3.71	1	0.011 3	.69	3.74	
Block 1	0.014	1				
Block 2	-0.014					
A-Wing Area	-8.33E-004	1	6.709E-003	-0.015	0.013	1.00
B-Wing Ratio	0.051	1	6.709E-003	0.037	0.065	1.00
C-Base Width	2.50E-003	1	6.709E-003	-0.012	0.017	1.00
D-Base Length	-0.061	1	6.709E-003	-0.075	-0.047	1.00
AB	-0.029	1	8.217E-003	-0.046	-0.011	1.00
AC	-0.037	1	8.217E-003	-0.055	-0.020	1.00
AD	0.044	1	8.217E-003	0.026	0.061	1.00
BC	0.046	1	8.217E-003	0.029	0.064	1.00
CD	-0.021	1	8.217E-003	-0.039	-3.830E-003	1.00
A2	-0.020	1	6.276E-003	-0.033	-6.800E-003	1.05
B2	-0.016	1	6.276E-003	-0.030	-3.050E-003	1.05
C2	-0.025	1	6.276E-003	-0.038	-0.012	1.05

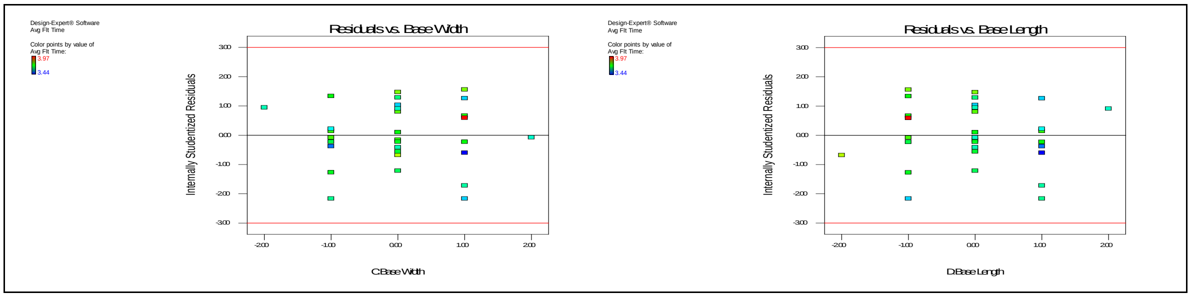
Final Equation in Terms of Coded Factors:

Avg Flt Time	=
+3.71	
-8.333E-004	* A
+0.051	* B
+2.500E-003	* C
-0.061	* D
-0.029	* A * B
-0.037	* A * C
+0.044	* A * D
+0.046	* B * C
-0.021	* C * D
-0.020	* A2
-0.016	* B2
-0.025	* C2

The residual analysis for the Average Flight Time response shows a little inequality of variance in Factor A.







For the Standard Deviation Flight Time response in the combined model: Only the CD interaction and the A^2 term are significant. A, C and D are added to maintain hierarchy.

Design Expert Output

Response 2 Stdev Flt Time

ANOVA for Response Surface Reduced Quadratic Model

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

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Block	5.339E-004	1	5.339E-004			
Model	5.890E-003	5	1.178E-003	2.85	0.0383	significant
A-Wing Area	*1.504E-005*	*1*	*1.504E-005*	*0.036*	*0.8505*	
C-Base Width	*2.042E-006*	*1*	*2.042E-006*	*4.93E-003*	*0.9446*	
D-Base Length	*6.934E-004*	*1*	*6.934E-004*	*1.68*	*0.2084*	
CD	*1.871E-003*	*1*	*1.871E-003*	*4.52*	*0.0445*	
A2	*3.309E-003*	*1*	*3.309E-003*	*7.99*	*0.0095*	
Residual	9.518E-003	23	4.138E-004			
Lack of Fit	*8.889E-003*	*19*	*4.679E-004*	*2.97*	*0.1499*	*not significant*
Pure Error	*6.292E-004*	*4*	*1.573E-004*			
Cor Total	0.016	29				

Std. Dev.	0.020	R-Squared	0.3822
Mean	0.069	Adj R-Squared	0.2479
C.V. %	29.48	Pred R-Squared	-0.3520
PRESS	0.021	Adeq Precision	6.331

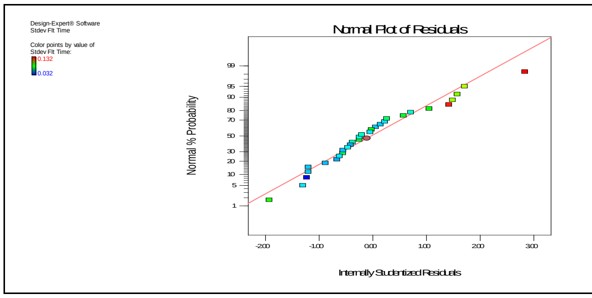
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	0.061	1	4.814E-003	0.052	0.071	
Block 1	-5.50E-003	1				
Block 2	5.50E-003					
A-Wing Area	-7.92E-004	1	4.153E-003	-9.382E-003	7.799E-003	1.00
C-Base Width	2.92E-004	1	4.153E-003	-8.299E-003	8.882E-003	1.00
D-Base Length	5.37E-003	1	4.153E-003	-3.215E-003	0.014	1.00
CD	-0.011	1	5.086E-003	-0.021	-2.917E-004	1.00
A2	0.011	1	3.814E-003	2.895E-003	0.019	1.01

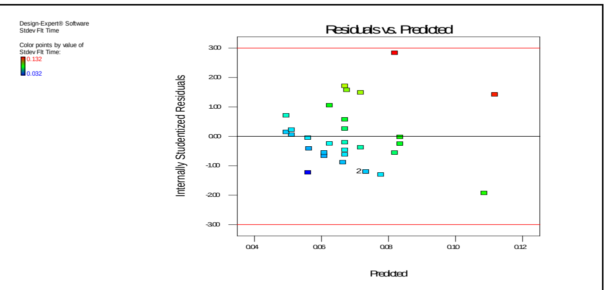
Final Equation in Terms of Coded Factors:

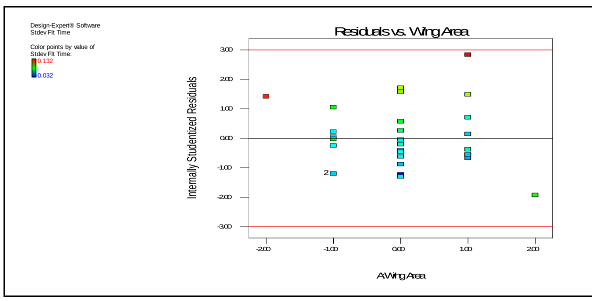
Stdev Flt Time =
+0.061

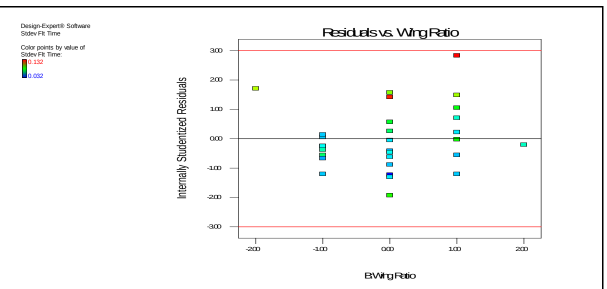
-7.917E-004	* A
+2.917E-004	* C
+5.375E-003	* D
-0.011	* C * D
+0.011	* A2

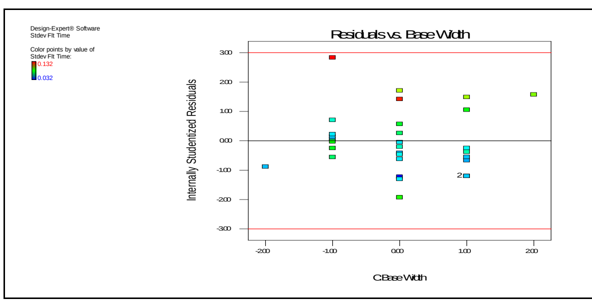
The residual analysis of the Standard Deviation Flight Time response shows in equality of variance in the Residuals vs. Predicted graph and in Factor B. Might want to try a transformation on the Standard Deviation Response.

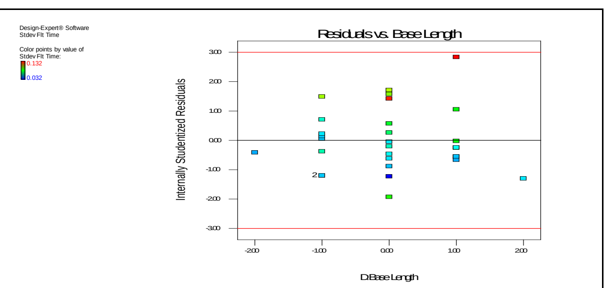




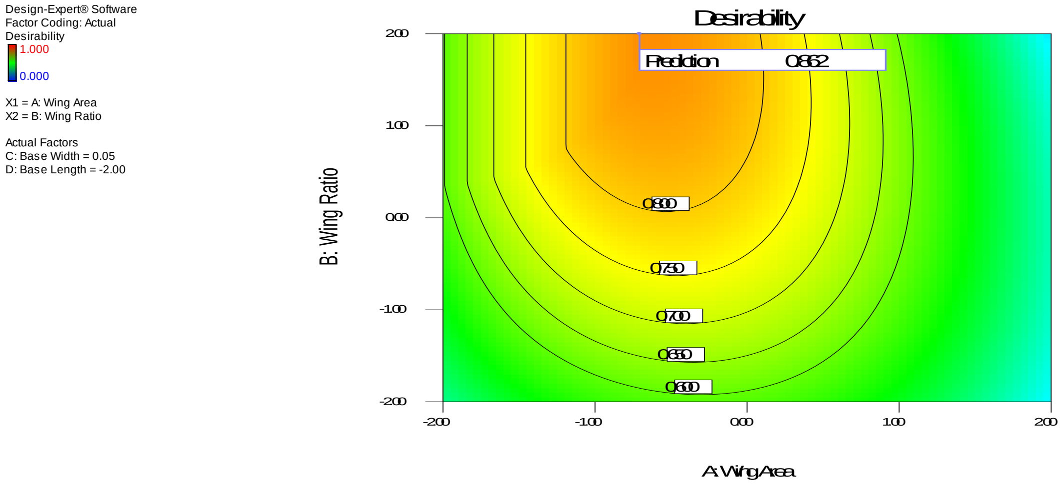


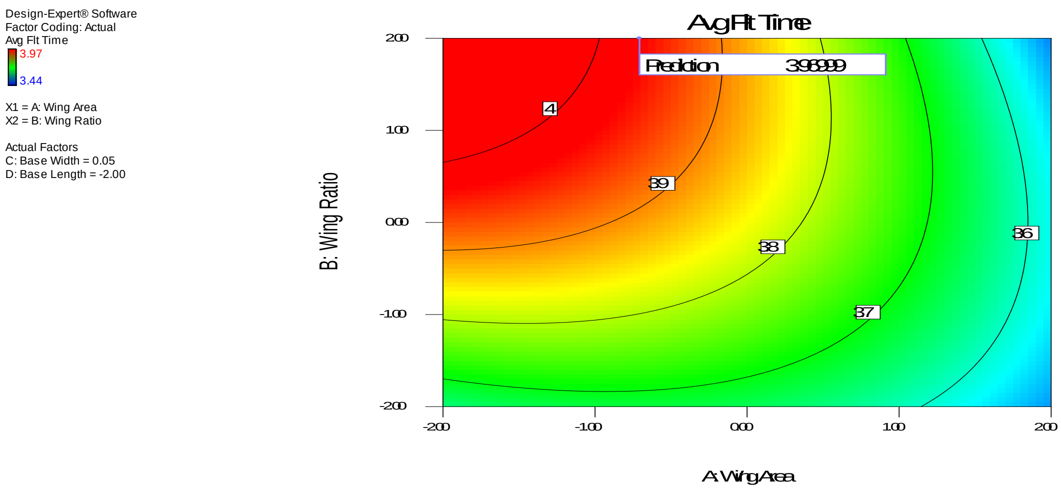






(d) What design would you recommend to maximize the flight time while simultaneously minimizing the standard deviation of flight time? A = -0.71, B = 2, C = 0.05, D = -2, in coded variables gives a predicted Flight Time of 3.967 and a predicted Standard Deviation of Flight Time of 0.058.





11.36. An article in the *Journal of Chromatography A* ("Optimization of the Capillary Electrophoresis Separation of Ranitidine and Related Compounds," Vol. 766, pp. 245-254) describes an experiment to optimize the production of ranitidine, a compound that is the primary active ingredient of Zantac, a pharmaceutical product used to treat ulcers, gastroesophageal reflux disease (a condition in which backward flow of acid from the stomach causes heartburn and injury of the esophagus), and other conditions where the stomach produces too much acid, such as Zollinger-Ellison syndrome. The authors used three factors (x_1 = pH of the buffer solution, x_2 = the electrophoresis voltage, and the concentration of one component of the buffer solution) in a central composite design. The response is chromatographic exponential function (CEF), which should be minimized. Table P11.12 shows the design.

P11.12 – The Ranitidine Separation Experiment

Std Ord	X1	X2	X3	CEF
1	-1	-1	-1	17.3
2	1	-1	-1	45.5
3	-1	1	-1	10.3
4	1	1	-1	11757.1
5	-1	-1	1	16.942
6	1	-1	1	25.4

7	-1	1	1	31697.2
8	1	1	1	12039.2
9	-1.68	0	0	7.5
10	1.68	0	0	6.3
11	0	-1.68	0	11.1
12	0	1.68	0	6.6664
13	0	0	-1.68	16548.7
14	0	0	1.68	26351.8
15	0	0	0	9.9
16	0	0	0	9.6
17	0	0	0	8.9
18	0	0	0	8.8
19	0	0	0	8.013
20	0	0	0	8.059

(a) Fit the second-order model to the CEF response. Analyze the residuals from this model. Does it seem that all model terms are necessary? There are many terms with low F-Values. The Adjusted R-Square is low relative to the R-Square indicating too many terms in the model.

Design Expert Output

Response 1 CEF
ANOVA for Response Surface Quadratic 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	1.480E+009	9	1.644E+008	5.66	0.0061	significant
A-pH	*4.543E+006*	*1*	*4.543E+006*	*0.16*	*0.7008*	
B-voltage	*2.247E+008*	*1*	*2.247E+008*	*7.73*	*0.0194*	
C-concen	*1.718E+008*	*1*	*1.718E+008*	*5.91*	*0.0353*	
AB	*7.896E+006*	*1*	*7.896E+006*	*0.27*	*0.6134*	
AC	*1.234E+008*	*1*	*1.234E+008*	*4.25*	*0.0662*	
BC	*1.279E+008*	*1*	*1.279E+008*	*4.40*	*0.0622*	
A2	*2.637E+005*	*1*	*2.637E+005*	*9.080E-003*	*0.9260*	
B2	*2.610E+005*	*1*	*2.610E+005*	*8.986E-003*	*0.9264*	
C2	*7.990E+008*	*1*	*7.990E+008*	*27.51*	*0.0004*	
Residual	2.905E+008	10	2.905E+007			
Lack of Fit	*2.905E+008*	*5*	*5.809E+007*	*9.710E+007*	*< 0.0001*	*significant*
Pure Error	*2.99*	*5*	*0.60*			
Cor Total	1.770E+009	19				

Std. Dev.	5389.41	R-Squared	0.8359
Mean	4929.71	Adj R-Squared	0.6882
C.V. %	109.32	Pred R-Squared	-0.3559
PRESS	2.40E+009	Adeq Precision	8.981

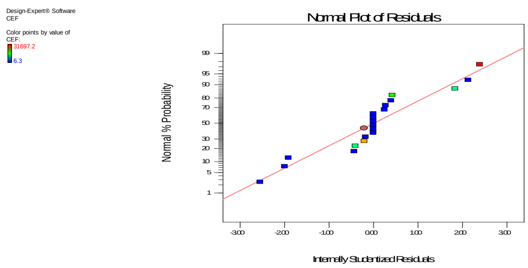
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	29.47	1	2198.06	-4868.12	4927.06	
A-pH	-576.75	1	1458.36	-3826.18	2672.69	1.00
B-voltage	4055.93	1	1458.36	806.49	7305.36	1.00
C-concentration	3546.59	1	1458.36	297.16	6796.03	1.00
AB	-993.48	1	1905.44	-5239.08	3252.11	1.00
AC	-3928.07	1	1905.44	-8173.66	317.53	1.00
BC	3998.68	1	1905.44	-246.91	8244.28	1.00
A2	-135.28	1	1419.68	-3298.51	3027.96	1.02

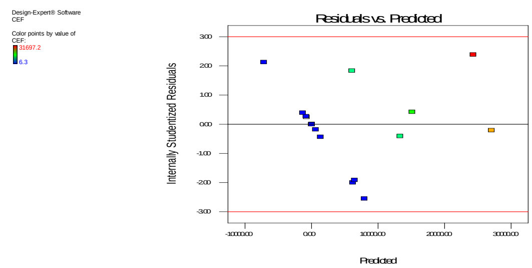
B2	-134.58	1	1419.68	-3297.81	3028.66	1.02
C2	7446.09	1	1419.68	4282.85	10609.33	1.02

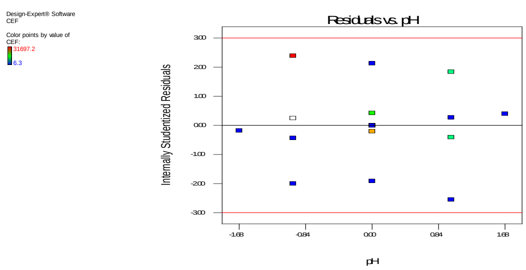
Final Equation in Terms of Coded Factors:

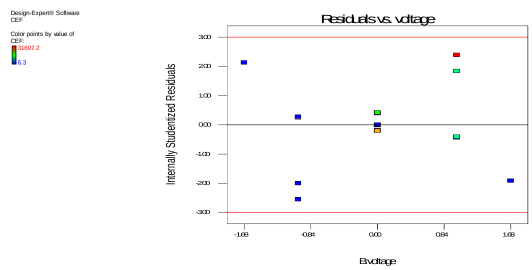
CEF =
+29.47
-576.75 * A
+4055.93 * B
+3546.59 * C
-993.48 * A * B
-3928.07 * A * C
+3998.68 * B * C
-135.28 * A2
-134.58 * B2
+7446.09 * C2

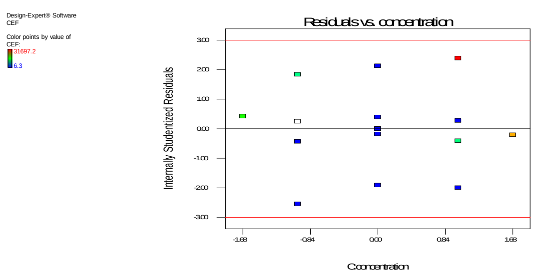
The Normal Plot of residuals is somewhat S-shaped.











(b) Reduce the model from part (a) as necessary. Did the model reduction improve the fit? The Adjusted R-Square improved. The only terms left in the model have significant F-Values.

Design Expert Output

Response 1 CEF
ANOVA for Response Surface Reduced Quadratic Model
Analysis of variance table [Partial sum of squares - Type III]

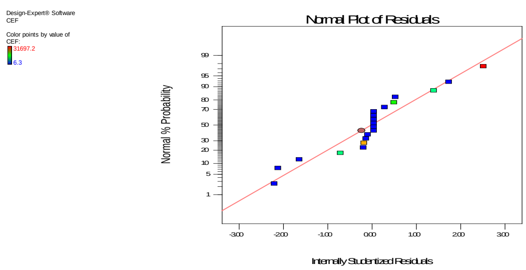
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	1.471E+009	6	2.452E+008	10.67	0.0002	significant

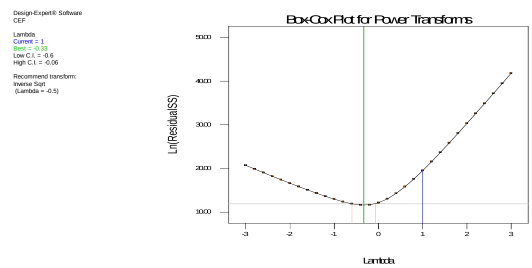
A-pH	*4.543E+006*	*1*	*4.543E+006*	*0.20*	*0.6640*	
B-voltage	*2.247E+008*	*1*	*2.247E+008*	*9.77*	*0.0080*	
C-concen	*1.718E+008*	*1*	*1.718E+008*	*7.47*	*0.0171*	
AC	*1.234E+008*	*1*	*1.234E+008*	*5.37*	*0.0374*	
BC	*1.279E+008*	*1*	*1.279E+008*	*5.56*	*0.0346*	
C2	*8.190E+008*	*1*	*8.190E+008*	*35.63*	*< 0.0001*	
Residual	2.988E+008	13	2.299E+007			
Lack of Fit	*2.988E+008*	*8*	*3.735E+007*	*6.244E+007*	*< 0.0001*	*significant*
Pure Error	*2.99*	*5*	*0.60*			
Cor Total	1.770E+009	19				

Std. Dev.	4794.47	R-Squared	0.8312
Mean	4929.71	Adj R-Squared	0.7533
C.V. %	97.26	Pred R-Squared	0.3793
PRESS	1.099E+009	Adeq Precision	11.957

(c) Does transformation of the CEF response seem like a useful idea? What aspect of either the data or the residuals suggests that transformation would be helpful?

The Normal Plot of Residuals still shows an S-shaped pattern in the reduced model. Suggest an inverse square root transformation.





(d) Fit a second-order model to the transformed CEF response. Analyze the residuals from this model. Does it seem that all model terms are necessary? What would you choose as the final model? Like the first model, there are too many terms. Many with low F-Values that could be deleted from the model.

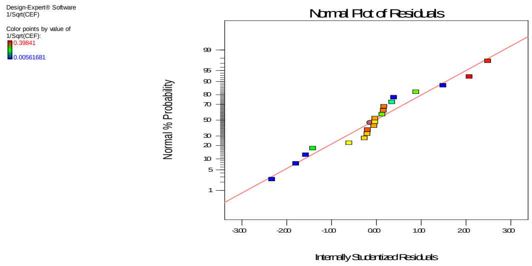
Design Expert Output

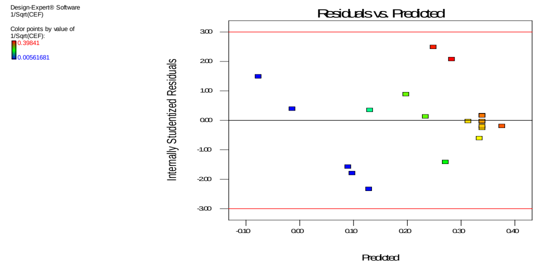
Response 1 CEF
Transform: Inverse Sqrt Constant: 0
ANOVA for Response Surface Quadratic Model
Analysis of variance table [Partial sum of squares - Type III]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	0.33	9	0.036	4.59	0.0130	significant
A-pH	*0.011*	*1*	*0.011*	*1.33*	*0.2757*	
B-voltage	*8.864E-003*	*1*	*8.864E-003*	*1.12*	*0.3155*	
C-concen	*4.802E-003*	*1*	*4.802E-003*	*0.60*	*0.4547*	
AB	*3.287E-003*	*1*	*3.287E-003*	*0.41*	*0.5344*	
AC	*0.016*	*1*	*0.016*	*1.97*	*0.1910*	
BC	*0.016*	*1*	*0.016*	*2.03*	*0.1850*	
A2	*1.819E-004*	*1*	*1.819E-004*	*0.023*	*0.8827*	
B2	*4.168E-003*	*1*	*4.168E-003*	*0.52*	*0.4853*	
C2	*0.27*	*1*	*0.27*	*33.61*	*0.0002*	
Residual	0.079	10	7.939E-003			
Lack of Fit	*0.078*	*5*	*0.016*	*73.24*	*0.0001*	*significant*
Pure Error	*1.069E-003*	*5*	*2.139E-004*			
Cor Total	0.41	19				

Std. Dev.	0.089	R-Squared	0.8050
Mean	0.23	Adj R-Squared	0.6295
C.V. %	38.33	Pred R-Squared	-0.4939
PRESS	0.61	Adeq Precision	7.192

The model residuals have improved.



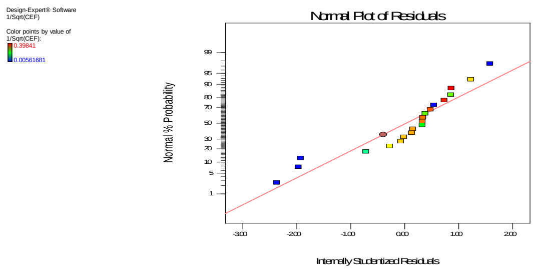


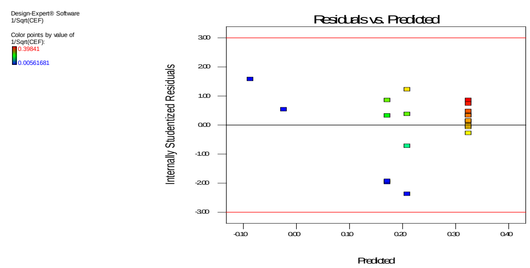
The reduced model includes only factor C and C^2.
Design Expert Output

Response 1 CEF
Transform: Inverse Sqrt Constant: 0
ANOVA for Response Surface Reduced Quadratic 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	0.27	2	0.13	16.57	0.0001	significant
C-conce	*4.802E-003*	*1*	*4.802E-003*	*0.59*	*0.4524*	
C2	*0.26*	*1*	*0.26*	*32.56*	*< 0.0001*	
Residual	0.14	17	8.119E-003			
Lack of Fit	*0.14*	*12*	*0.011*	*53.36*	*0.0002*	*significant*
Pure Error	*1.069E-003*	*5*	*2.139E-004*			
Cor Total	0.41	19				

Std. Dev.	0.090	R-Squared	0.6610
Mean	0.23	Adj R-Squared	0.6211
C.V. %	38.76	Pred R-Squared	0.4588
PRESS	0.22	Adeq Precision	11.782





(e) Suppose you had some information that suggests that the separation process malfunctioned during run 7. Delete this run and analyze the data from this experiment again. Deleting run 7 does not have much impact on the experiment.

Design Expert Output

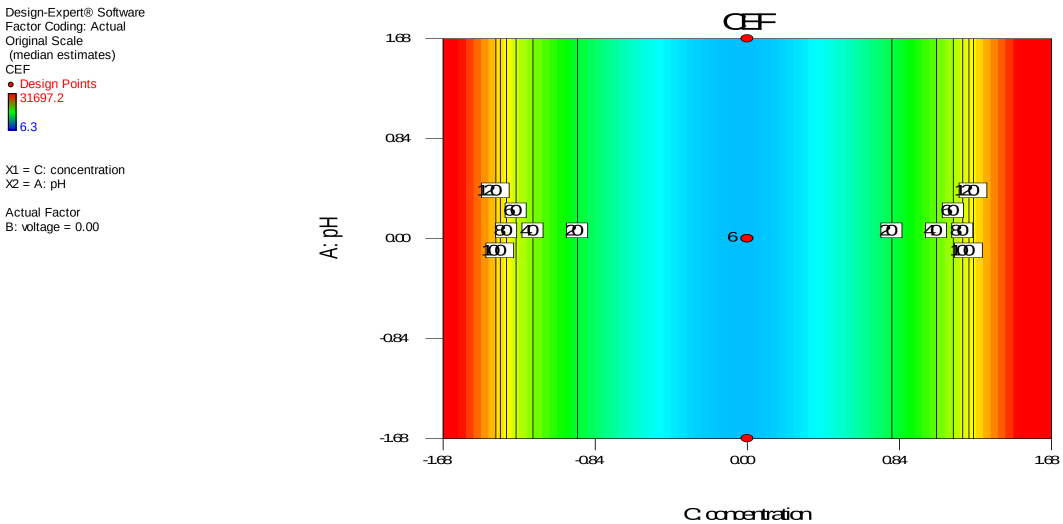
Response 1 CEF
Transform: Inverse Sqrt Constant: 0
ANOVA for Response Surface Reduced Quadratic Model
Analysis of variance table [Partial sum of squares - Type III]

Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Model	0.25	2	0.12	18.51	< 0.0001	significant

C-conce	*2.928E-004*	*1*	*2.928E-004*	*0.044*	*0.8366*	
C2	*0.25*	*1*	*0.25*	*37.01*	*< 0.0001*	
Residual	0.11	16	6.658E-003			
Lack of Fit	*0.11*	*11*	*9.587E-003*	*44.83*	*0.0003*	*significant*
Pure Error	*1.069E-003*	*5*	*2.139E-004*			
Cor Total	0.35	18				

Std. Dev.	0.082	R-Squared	0.6982
Mean	0.24	Adj R-Squared	0.6605
C.V. %	33.39	Pred R-Squared	0.5392
PRESS	0.16	Adeq Precision	11.599

(f) What conditions would you recommend to minimize CEF? Run C at the zero level. CEF is minimized.



11.37. An article in the *Electronic Journal of Biotechnology* ("Optimization of Medium Composition for Transglutaminase Production by a Brazilian Soil *Streptomyces* sp," available at http://www.ejbiotechnology.info/content/vol10/issue4/full/10.index.html) describes the use of designed experiments to improve the medium for cells used in a new microbial source of transglutaminase (MTGase), an enzyme that catalyzes an acyl transfer reaction using peptide-bond glutamine residues as acyl donors and some primary amines as acceptors. Reactions catalyzed by MTGase can be used in food processing. The article describes two phases of experimentation – screening with a fractional factorial and optimization. We will use only the optimization experiment. The design was a central composite design in four factors – x_1 = KH_2PO_4, x_2 = MgSO4-7H_2O, x_3 = soybean flower, and x_4 = peptone. MTGase activity is the response, which should be maximized. Table P11.13 contains the design and the response data.

Table P11.13 – The MTGase Optimization Experiment

Std Ord	X1	X2	X3	X4	MTGase Activity
1	-1	-1	-1	-1	0.87
2	1	-1	-1	-1	0.74
3	-1	1	-1	-1	0.51
4	1	1	-1	-1	0.99
5	-1	-1	1	-1	0.67

6	1	-1	1	-1	0.72
7	-1	1	1	-1	0.81
8	1	1	1	-1	1.01
9	-1	-1	-1	1	1.33
10	1	-1	-1	1	0.7
11	-1	1	-1	1	0.82
12	1	1	-1	1	0.78
13	-1	-1	1	1	0.36
14	1	-1	1	1	0.23
15	-1	1	1	1	0.21
16	1	1	1	1	0.44
17	-2	0	0	0	0.56
18	2	0	0	0	0.49
19	0	-2	0	0	0.57
20	0	2	0	0	0.81
21	0	0	-2	0	0.9
22	0	0	2	0	0.65
23	0	0	0	-2	0.91
24	0	0	0	2	0.49
25	0	0	0	0	1.43
26	0	0	0	0	1.17
27	0	0	0	0	1.5

(a) Fit a second-order model to the MSGase activity response. The second-order model includes All 4 quadratic terms – A^2, B^2, C^2 and D^2; the two-factor interactions – AB and CD (AD is borderline and kept in the model) and the main effects C and D. A and B were left in the model for model hierarchy.

Design Expert Output

Response 1 MTGase

ANOVA for Response Surface Reduced Quadratic 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	2.49	11	0.23	10.29	< 0.0001	significant
A-X1	*5.042E-004*	*1*	*5.042E-004*	*0.023*	*0.8817*	
B-X2	*7.704E-003*	*1*	*7.704E-003*	*0.35*	*0.5630*	
C-X3	*0.32*	*1*	*0.32*	*14.73*	*0.0016*	
D-X4	*0.22*	*1*	*0.22*	*9.92*	*0.0066*	
AB	*0.18*	*1*	*0.18*	*8.30*	*0.0114*	
AD	*0.086*	*1*	*0.086*	*3.89*	*0.0674*	
CD	*0.39*	*1*	*0.39*	*17.60*	*0.0008*	
A2	*0.90*	*1*	*0.90*	*41.09*	*< 0.0001*	
B2	*0.58*	*1*	*0.58*	*26.28*	*0.0001*	
C2	*0.44*	*1*	*0.44*	*19.94*	*0.0005*	
D2	*0.56*	*1*	*0.56*	*25.49*	*0.0001*	
Residual	0.33	15	0.022			
Lack of Fit	*0.27*	*13*	*0.021*	*0.69*	*0.7314*	*not significant*
Pure Error	*0.060*	*2*	*0.030*			
Cor Total	2.82	26				

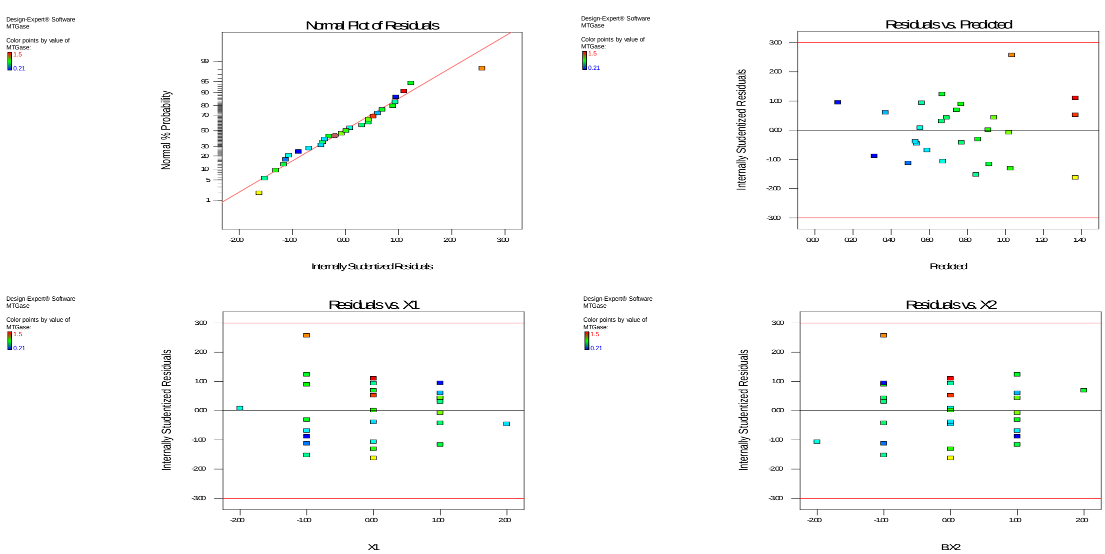
Std. Dev.	0.15	R-Squared	0.8830
Mean	0.77	Adj R-Squared	0.7972
C.V. %	19.38	Pred R-Squared	0.6452
PRESS	1.00	Adeq Precision	12.597

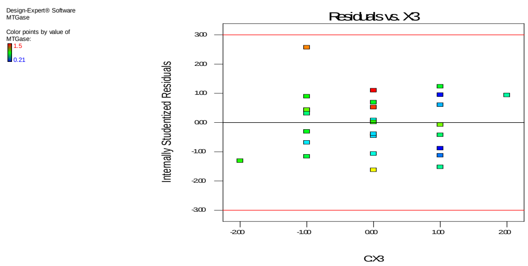
Factor	**Coefficient Estimate**	**df**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	1.37	1	0.086	1.18	1.55	
A-X1	-4.583E-003	1	0.030	-0.069	0.060	1.00
B-X2	0.018	1	0.030	-0.047	0.082	1.00
C-X3	-0.12	1	0.030	-0.18	-0.052	1.00
D-X4	-0.095	1	0.030	-0.16	-0.031	1.00
AB	0.11	1	0.037	0.028	0.19	1.00
AD	-0.073	1	0.037	-0.15	5.940E-003	1.00
CD	-0.16	1	0.037	-0.23	-0.077	1.00
A2	-0.21	1	0.032	-0.27	-0.14	1.25
B2	-0.16	1	0.032	-0.23	-0.096	1.25
C2	-0.14	1	0.032	-0.21	-0.075	1.25
D2	-0.16	1	0.032	-0.23	-0.094	1.25

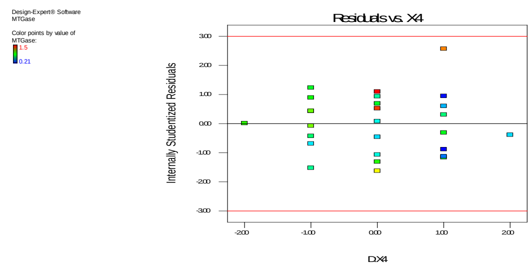
Final Equation in Terms of Actual Factors:

MTGase =	
+1.36667	
-4.58333E-003	* X1
+0.017917	* X2
-0.11625	* X3
-0.095417	* X4
+0.10688	* X1 * X2
-0.073125	* X1 * X4
-0.15563	* X3 * X4
-0.20594	* X12
-0.16469	* X22
-0.14344	* X32
-0.16219	* X42

(b) Analyze the residuals from this model. Run #9 looks to be an outlier. The residual plots look good for model adequacy.







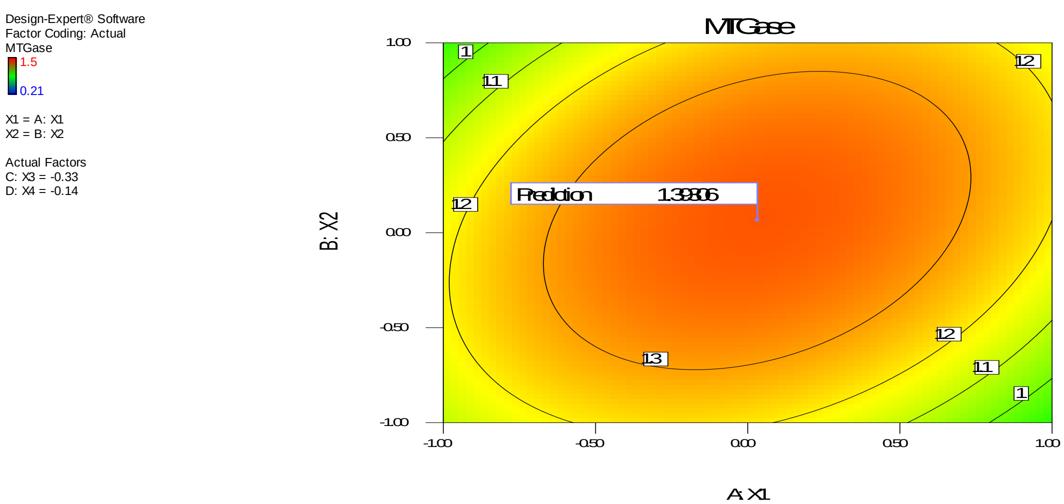
(c) Recommend operating conditions that maximize MTGase activity. MTGase can be maximized with X1 = 0.03, X2 = 0.07, X3 = -0.33 and X4 = -0.14. The MTGase response will be 1.39,

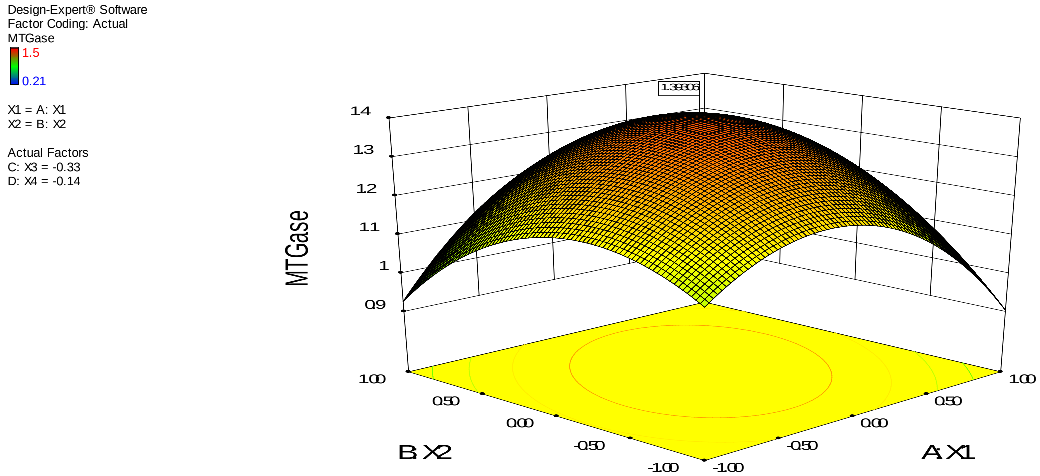
Constraints

		Lower	Upper	Lower	Upper	
Name	**Goal**	**Limit**	**Limit**	**Weight**	**Weight**	**Importance**
A:X1	is in range	-2	2	1	1	3
B:X2	is in range	-2	2	1	1	3
C:X3	is in range	-2	2	1	1	3
D:X4	is in range	-2	2	1	1	3
MTGase	maximize	0.21	1.5	1	1	3

Solutions

Number	**X1**	**X2**	**X3**	**X4**	**MTGase**	**Desirability**	
1	0.03	0.07	-0.33	-0.14	1.39306	0.917	Selected



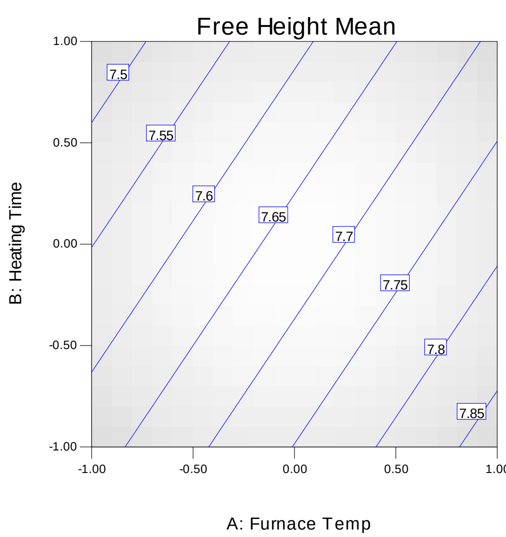

Design-Expert® Software
Factor Coding: Actual
MTGase
1.5
0.21
X1 = A: X1
X2 = B: X2
Actual Factors
C: X3 = -0.33
D: X4 = -0.14
MTGase
B: X2
A: X1

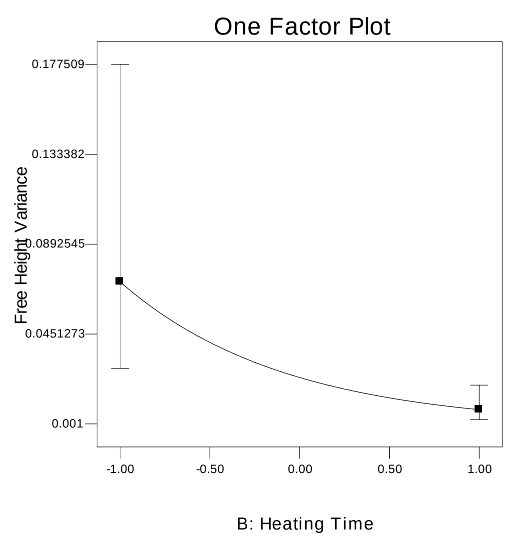
Chapter 12
Robust Parameter Design and Process Robustness Studies
Solutions

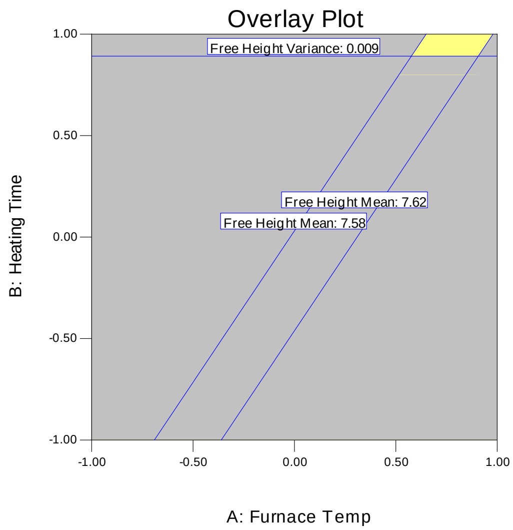
12.1. Reconsider the leaf spring experiment in Table 12.1. Suppose that the objective is to find a set of conditions where the mean free height is as close as possible to 7.6 inches with a variance of free height as small as possible. What conditions would you recommend to achieve these objectives?

A	*B*	*C*	*D*	*E*(-)	*E*(+)	$\bar{y}$	s^2
-	-	-	-	7.78,7.78, 7.81	7.50, 7,25, 7.12	7.54	0.090
+	-	-	+	8.15, 8.18, 7.88	7.88, 7.88, 7.44	7.90	0.071
-	+	-	+	7.50, 7.56, 7.50	7.50, 7.56, 7.50	7.52	0.001
+	+	-	-	7.59, 7.56, 7.75	7.63, 7.75, 7.56	7.64	0.008
-	-	+	+	7.54, 8.00, 7.88	7.32, 7.44, 7.44	7.60	0.074
+	-	+	-	7.69, 8.09, 8.06	7.56, 7.69, 7.62	7.79	0.053
-	+	+	-	7.56, 7.52, 7.44	7.18, 7.18, 7.25	7.36	0.030
+	+	+	+	7.56, 7.81, 7.69	7.81, 7.50, 7.59	7.66	0.017

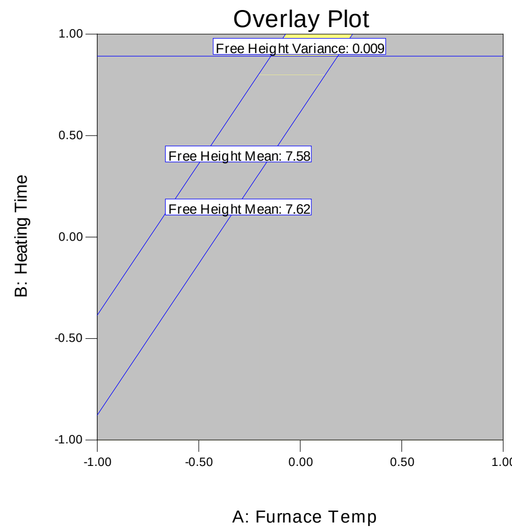
By overlaying the contour plots for Free Height Mean and the Free Height Variance, optimal solutions can be found. To minimize the variance, factor B must be at the high level while factors *A* and *D* are adjusted to assure a mean of 7.6. The two overlay plots below set factor *D* at both low and high levels. Therefore, a mean as close as possible to 7.6 with minimum variance of 0.008 can be achieved at *A* = 0.78, *B* = +1, and *D* = -1. This can also be achieved with *A* = +0.07, *B* = +1, and *D* = +1.







Factor D = -1



Factor D = +1

12.2. Consider the bottle filling experiment in Problem 6-18. Suppose that the percentage of carbonation (*A*) is a noise variable ($\sigma_z^2 = 1$ in coded units).

(a) Fit the response model to these data. Is there a robust design problem?

The following is the analysis of variance with all terms in the model followed by a reduced model. Because the noise factor *A* is significant, and the AB interaction is moderately significant, there is a robust design problem.

Design Expert Output

Response: Fill Deviation
ANOVA for Response Surface Reduced Cubic Model
Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Cor Total	300.05	3				
Model	73.00	7	10.43	16.69	0.0003	significant
A	*36.00*	*1*	*36.00*	*57.60*	*< 0.0001*	
B	*20.25*	*1*	*20.25*	*32.40*	*0.0005*	
C	*12.25*	*1*	*12.25*	*19.60*	*0.0022*	
AB	*2.25*	*1*	*2.25*	*3.60*	*0.0943*	
AC	*0.25*	*1*	*0.25*	*0.40*	*0.5447*	
BC	*1.00*	*1*	*1.00*	*1.60*	*0.2415*	
ABC	*1.00*	*1*	*1.00*	*1.60*	*0.2415*	
Pure Error	5.00	8	0.63			
Cor Total	78.00	15				

Based on the above analysis, the AC, BC, and ABC interactions are removed from the model and used as error.

Design Expert Output

Response: Fill Deviation
ANOVA for Response Surface Reduced Cubic 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
A	*36.00*	*1*	*36.00*	*54.62*	*< 0.0001*	

B	*20.25*	*1*	*20.25*	*30.72*	*0.0002*	
C	*12.25*	*1*	*12.25*	*18.59*	*0.0012*	
AB	*2.25*	*1*	*2.25*	*3.41*	*0.0917*	
Residual	7.25	11	0.66			
Lack of Fit	2.25	3	0.75	1.20	0.3700	not significant
Pure Error	5.00	8	0.63			
Cor Total	78.00	15				

The Model F-value of 26.84 implies there is a 0.01% chance that a "Model F-Value" this large could occur due to noise.

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

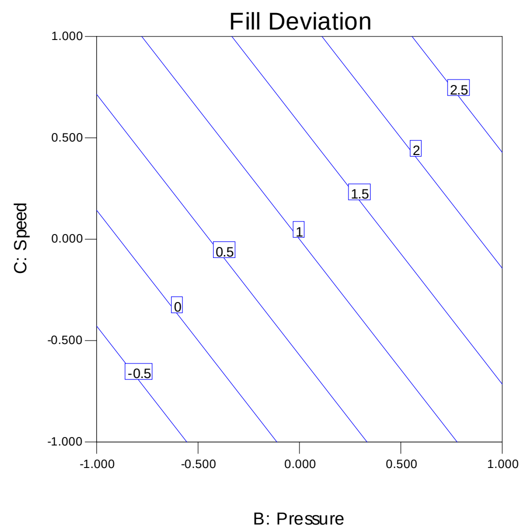
Final Equation in Terms of Coded Factors:

Fill Deviation =
+1.00
+1.50 * A
+1.13 * B
+0.88 * C
+0.38 * A * B

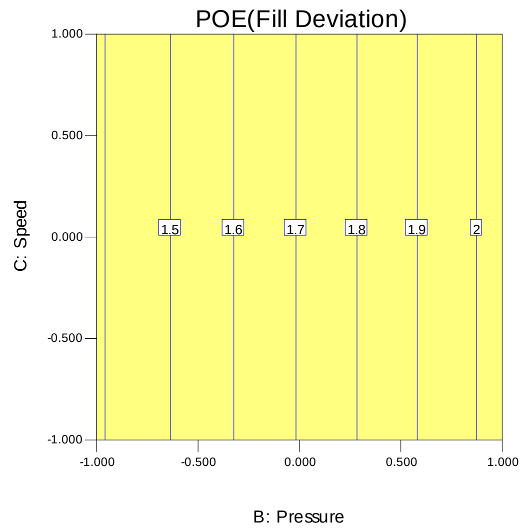
(b) Find the mean model and either the variance model or the POE.

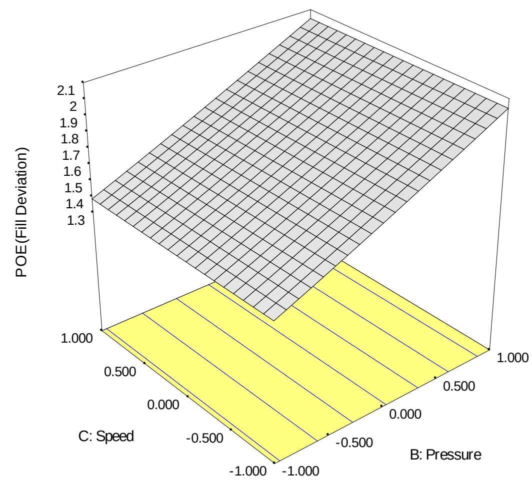
From the final equation shown in the above analysis, the mean model and corresponding contour plot is shown below.

$$E_z\left[y\left(\mathbf{x},z_1\right)\right]=1+1.13x_2+0.88x_3$$



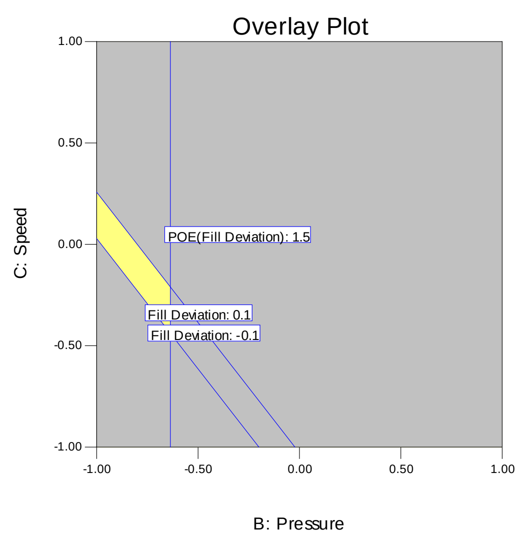
Contour and 3-D plots of the POE are shown below.





(c) Find a set of conditions that result in mean fill deviation as close to zero as possible with minimum transmitted variance.

The overlay plot below identifies a an operating region for pressure and speed that in a mean fill deviation as close to zero as possible with minimum transmitted variance.



12.3. Consider the experiment in Problem 11-12. Suppose that temperature is a noise variable ($\sigma_z^2 = 1$ in coded units). Fit response models for both responses. Is there a robust design problem with respect to both responses? Find a set of conditions that maximize conversion with activity between 55 and 60 and that minimize variability transmitted from temperature.

The analysis and models as found in problem 11-12 are shown below for both responses. There is a robust design problem with regards to the conversion response because of the significance of factor *B*, temperature, and the *BC* interaction. However, temperature is not significant in the analysis of the second response, activity.

Design Expert Output

Response: Conversion

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	2555.73	9	283.97	12.76	0.0002	significant
A	*14.44*	*1*	*14.44*	*0.65*	*0.4391*	
B	*222.96*	*1*	*222.96*	*10.02*	*0.0101*	
C	*525.64*	*1*	*525.64*	*23.63*	*0.0007*	
A^2	*48.47*	*1*	*48.47*	*2.18*	*0.1707*	
B^2	*124.48*	*1*	*124.48*	*5.60*	*0.0396*	
C^2	*388.59*	*1*	*388.59*	*17.47*	*0.0019*	
AB	*36.13*	*1*	*36.13*	*1.62*	*0.2314*	
AC	*1035.13*	*1*	*1035.13*	*46.53*	*< 0.0001*	
BC	*120.12*	*1*	*120.12*	*5.40*	*0.0425*	
Residual	222.47	10	22.25			
Lack of Fit	*56.47*	*5*	*11.29*	*0.34*	*0.8692*	*not significant*
Pure Error	*166.00*	*5*	*33.20*			
Cor Total	287.28	19				

The Model F-value of 12.76 implies the model is significant. There is only a 0.02% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	4.72	R-Squared	0.9199
Mean	78.30	Adj R-Squared	0.8479
C.V.	6.02	Pred R-Squared	0.7566
PRESS	676.22	Adeq Precision	14.239

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	81.09	1	1.92	76.81	85.38	
A-Time	1.03	1	1.28	-1.82	3.87	1.00
B-Temperature	4.04	1	1.28	1.20	6.88	1.00
C-Catalyst	6.20	1	1.28	3.36	9.05	1.00
A2	-1.83	1	1.24	-4.60	0.93	1.02
B2	2.94	1	1.24	0.17	5.71	1.02
C2	-5.19	1	1.24	-7.96	-2.42	1.02
AB	2.13	1	1.67	-1.59	5.84	1.00
AC	11.38	1	1.67	7.66	15.09	1.00
BC	-3.87	1	1.67	-7.59	-0.16	1.00

Final Equation in Terms of Coded Factors:

Conversion	=
+81.09	
+1.03	* A
+4.04	* B
+6.20	* C
-1.83	* A2
+2.94	* B2
-5.19	* C2
+2.13	* A * B
+11.38	* A * C
-3.87	* B * C

Design Expert Output

Response: Activity

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	253.20	3	84.40	39.63	< 0.0001	significant
A	*175.35*	*1*	*175.35*	*82.34*	*< 0.0001*	
C	*67.91*	*1*	*67.91*	*31.89*	*< 0.0001*	
A^2	*9.94*	*1*	*9.94*	*4.67*	*0.0463*	
Residual	34.07	16	2.13			
Lack of Fit	*30.42*	*11*	*2.77*	*3.78*	*0.0766*	*not significant*
Pure Error	*3.65*	*5*	*0.73*			

Cor Total	287.28	19

The Model F-value of 39.63 implies the model is significant. There is only a 0.01% chance that a "Model F-Value" this large could occur due to noise.

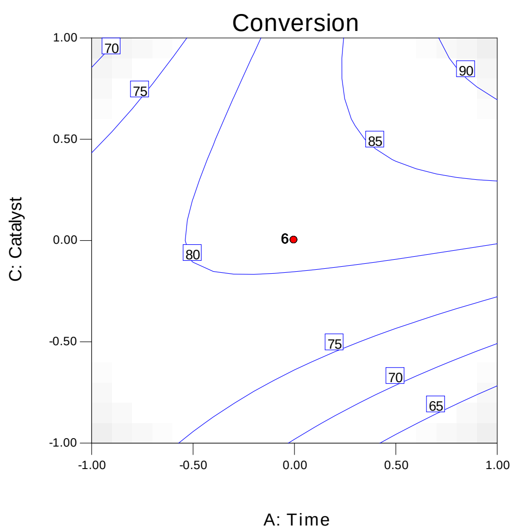
Std. Dev.	1.46	R-Squared	0.8814
Mean	60.51	Adj R-Squared	0.8591
C.V.	2.41	Pred R-Squared	0.6302
PRESS	106.24	Adeq Precision	20.447

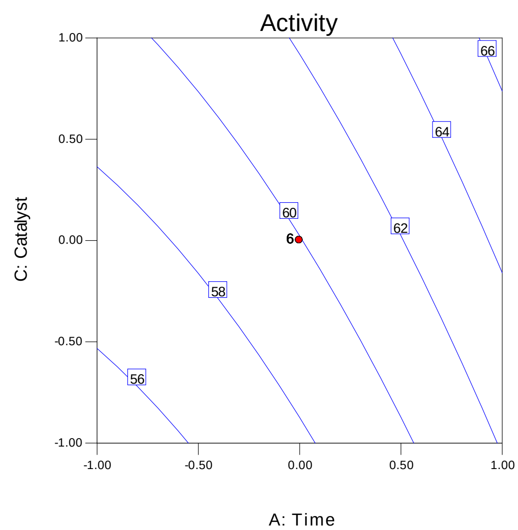
Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	59.95	1	0.42	59.06	60.83	
A-Time	3.58	1	0.39	2.75	4.42	1.00
C-Catalyst	2.23	1	0.39	1.39	3.07	1.00
A^2	0.82	1	0.38	0.015	1.63	1.00

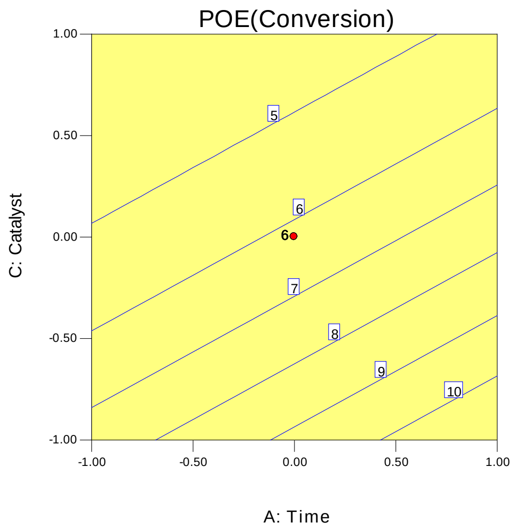
Final Equation in Terms of Coded Factors:

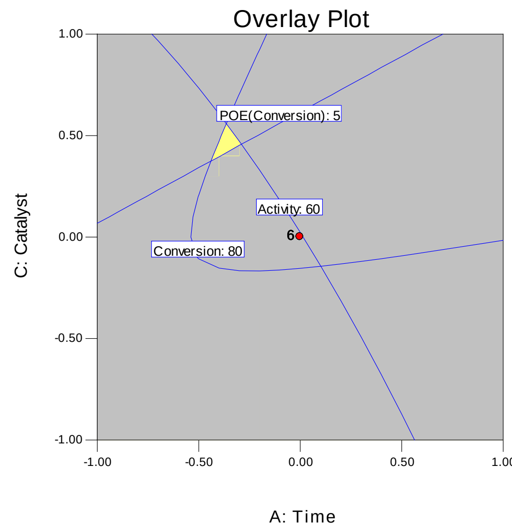
Activity =
+59.95
+3.58 * A
+2.23 * C
+0.82 * A2

The following contour plots of conversion, activity, and POE and the corresponding optimization plot identify a region where conversion is maximized, activity is between 55 and 60, and the transmitted variability from temperature is minimized. Factor *A* is set at 0.5 while *C* is set at 0.4.









12.4. Reconsider the leaf spring experiment from Table 12.1. Suppose that factors *A*, *B* and *C* are controllable variables, and that factors *D* and *E* are noise factors. Set up a crossed array design to investigate this problem, assuming that all of the two-factor interactions involving the controllable variables are thought to be important. What type of design have you obtained?

The following experimental design has a 2^3 inner array for the controllable variables and a 2^2 outer array for the noise factors. A total of 32 runs are required.

			Outer Array				
Inner Array			*D*	-1	1	-1	1
A	*B*	*C*	*E*	-1	-1	1	1
-1	-1	-1					
1	-1	-1					
-1	1	-1					
1	1	-1					
-1	-1	1					
1	-1	1					
-1	1	1					
1	1	1					

12.5. Continuation of Problem 12.4. Reconsider the leaf spring experiment from Table 12.1. Suppose that *A*, *B* and *C* are controllable factors and that factors *D* and *E* are noise factors. Show how a combined array design can be employed to investigate this problem that allows all two-factor interactions to be estimated and only requires 16 runs. Compare this with the crossed array design from Problem 12.4. Can you see how in general combined array designs have fewer runs than crossed array designs?

The following experiment is a 2^{5-1} fractional factorial experiment where the controllable factors are *A*, *B*, and *C* and the noise factors are *D* and *E*. Only 16 runs are required versus the 32 runs required for the crossed array design in problem 12.4.

A	*B*	*C*	*D*	*E*	Free Height
-	-	-	-	+	
+	-	-	-	-	
-	+	-	-	-	
+	+	-	-	+	
-	-	+	-	-	
+	-	+	-	+	
-	+	+	-	+	
+	+	+	-	-	
-	-	-	+	-	
+	-	-	+	+	
-	+	-	+	+	
+	+	-	+	-	
-	-	+	+	+	
+	-	+	+	-	
-	+	+	+	-	
+	+	+	+	+	

12.6. Consider the connector pull-off force experiment shown in Table 12.2. What main effects and interactions involving the controllable variables can be estimated with this design? Remember that all of the controllable variables are quantitative factors.

The design in Table 12.2 contains a 3^{4-2} inner array for the controllable variables. This is a resolution III design which aliases the main effects with two factor interactions. The alias table below identifies the alias structure for this design. Because of the partial aliasing in this design, it is difficult to interpret the interactions.

Design Expert Output

Alias Matrix

[Est. Terms]	**Aliased Terms**
[Intercept]	= Intercept - BC - BD - CD
[A]	= A - 0.5 * BC - 0.5 * BD - 0.5 * CD
[B]	= B - 0.5 * AC - 0.5 * AD
[C]	= C - 0.5 * AB - 0.5 * AD
[D]	= D - 0.5 * AB - 0.5 * AC
[A2]	= A2 + 0.5 * BC + 0.5 * BD + 0.5 * CD
[B2]	= B2 + 0.5 * AC - 0.5 * AD + CD
[C2]	= C2 - 0.5 * AB + 0.5 * AD + BD
[D2]	= D2 + 0.5 * AB - 0.5 * AC + BC

12.7. Consider the connector pull-off force experiment shown in Table 12.2. Show how an experiment can be designed for this problem that will allow a full quadratic model to be fit in the controllable variables along all main effects of the noise variables and their interactions with the controllable variables. How many runs will be required in this design? How does this compare with the design in Table 12.2?

There are several designs that can be employed to achieve the requirements stated above. Below is a small central composite design with the axial points removed for the noise variables. Five center points are also included which brings the total runs to 35. As shown in the alias analysis, the full quadratic model for the controllable variables is achieved.

A	*B*	*C*	*D*	*E*	*F*	*G*
+1	+1	+1	-1	+1	+1	+1
+1	+1	-1	+1	-1	+1	-1
+1	+1	-1	+1	+1	-1	+1
+1	-1	+1	+1	-1	+1	+1
-1	+1	+1	-1	-1	+1	-1
+1	-1	-1	-1	+1	-1	-1
-1	+1	-1	+1	+1	-1	+1
+1	+1	+1	+1	-1	+1	-1
+1	-1	+1	-1	-1	-1	+1
-1	-1	-1	-1	+1	+1	-1
-1	+1	-1	+1	-1	-1	-1
+1	+1	+1	-1	+1	-1	-1
+1	-1	-1	+1	-1	-1	-1
-1	-1	+1	-1	-1	-1	+1
-1	+1	-1	-1	-1	+1	+1
+1	-1	-1	-1	-1	+1	+1
-1	+1	-1	-1	+1	+1	+1
+1	-1	-1	+1	+1	+1	+1
-1	-1	+1	+1	+1	-1	+1
-1	-1	+1	+1	+1	+1	-1
-1	+1	+1	+1	-1	-1	+1
-1	-1	-1	-1	-1	-1	-1
-2.17	0	0	0	0	0	0
2.17	0	0	0	0	0	0
0	-2.17	0	0	0	0	0
0	2.17	0	0	0	0	0
0	0	-2.17	0	0	0	0
0	0	2.17	0	0	0	0
0	0	0	-2.17	0	0	0
0	0	0	2.17	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

Design Expert Output

Alias Matrix

[Est. Terms]	**Aliased Terms**
[Intercept]	= Intercept
[A]	= A
[B]	= B
[C]	= C
[D]	= D
[E]	= E + 0.211 * EG + 0.789 * FG
[F]	= F - EF - EG
[G]	= G - EF - 0.158 * EG + 0.158 * FG
[A^2]	= A^2
[B^2]	= B^2
[C^2]	= C^2
[D^2]	= D^2
[E^2]	= $E^2 + F^2 + G^2$
[AB]	= AB - 0.105 * EG - 0.895 * FG
[AC]	= AC - 0.158 * EG + 0.158 * FG
[AD]	= AD + 0.421 * EG + 0.579 * FG
[AE]	= AE - 0.474 * EG + 0.474 * FG
[AF]	= AF + EF + 1.05 * EG - 0.0526 * FG
[AG]	= AG + EF + 1.05 * EG - 0.0526 * FG
[BC]	= BC - 0.263 * EG + 0.263 * FG
[BD]	= BD - EF - 0.158 * EG + 0.158 * FG
[BE]	= BE - 0.368 * EG + 0.368 * FG
[BF]	= BF + 1.11 * EG - 0.105 * FG
[BG]	= BG + EF + 0.421 * EG - 0.421 * FG
[CD]	= CD - 0.421 * EG + 0.421 * FG
[CE]	= CE - EF + 0.158 * EG + 0.842 * FG
[CF]	= CF - EF - 0.211 * EG + 0.211 * FG
[CG]	= CG - 1.21 * EG + 0.211 * FG

[DE]	= DE - 0.842 * EG - 0.158 * FG
[DF]	= DF - 0.211 * EG + 0.211 * FG
[DG]	= DG - EF + 0.263 * EG - 0.263 * FG

12.8. Consider the experiment in Problem 11-11. Suppose that pressure is a noise variable ($\sigma_z^2 = 1$ in coded units). Fit the response model for the viscosity response. Find a set of conditions that result in viscosity as close as possible to 600 and that minimize the variability transmitted from the noise variable pressure.

Design Expert Output

Response: Viscosity

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	85467.33	6	14244.56	12.12	0.0012	significant
A	*703.12*	*1*	*703.12*	*0.60*	*0.4615*	
B	*6105.12*	*1*	*6105.12*	*5.19*	*0.0522*	
C	*5408.00*	*1*	*5408.00*	*4.60*	*0.0643*	
A2	*21736.93*	*1*	*21736.93*	*18.49*	*0.0026*	
C2	*5153.80*	*1*	*5153.80*	*4.38*	*0.0696*	
AC	*47742.25*	*1*	*47742.25*	*40.61*	*0.0002*	
Residual	9404.00	8	1175.50			
Lack of Fit	*7922.00*	*6*	*1320.33*	*1.78*	*0.4022*	*not significant*
Pure Error	*1482.00*	*2*	*741.00*			
Cor Total	94871.33	14				

The Model F-value of 12.12 implies the model is significant. There is only a 0.12% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	34.29	R-Squared	0.9009
Mean	575.33	Adj R-Squared	0.8265
C.V.	5.96	Pred R-Squared	0.6279
PRESS	35301.77	Adeq Precision	11.731

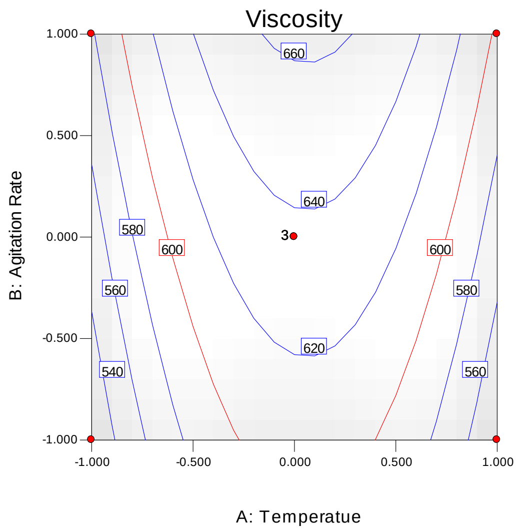
Final Equation in Terms of Coded Factors:

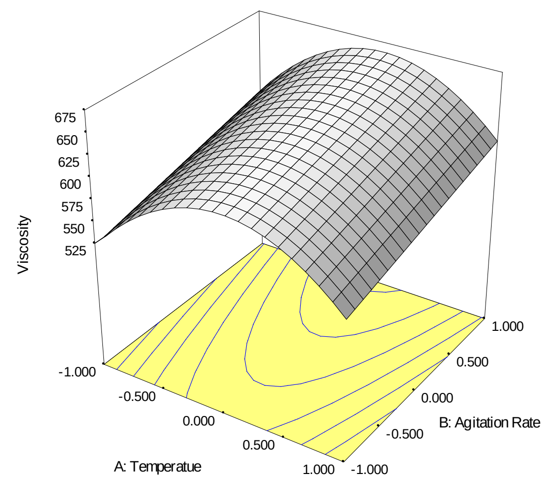
Viscosity	=
+636.00	
+9.37	* A
+27.62	* B
-26.00	* C
-76.50	* A2
-37.25	* C2
+109.25	* A * C

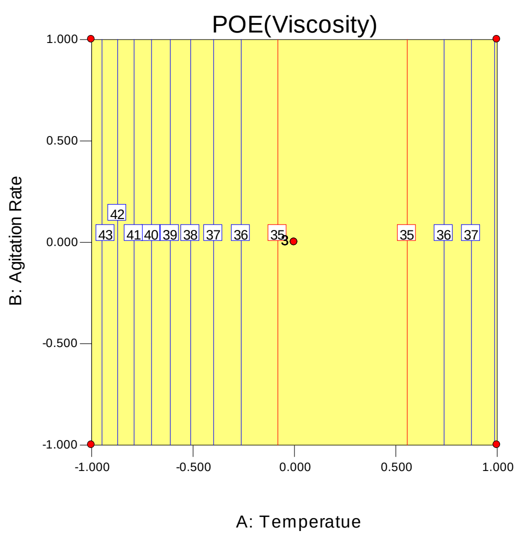
From the final equation shown in the above analysis, the mean model is shown below.

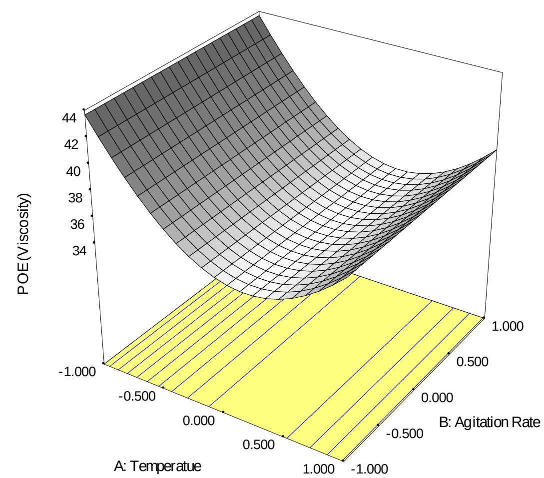
$$E_z\left[y(\mathbf{x}, z_1)\right] = 636.00 + 9.37x_1 + 27.62x_2 - 26.00x_3 - 76.50x_1^2 - 37.25x_3^2 + 109.25x_1x_3$$

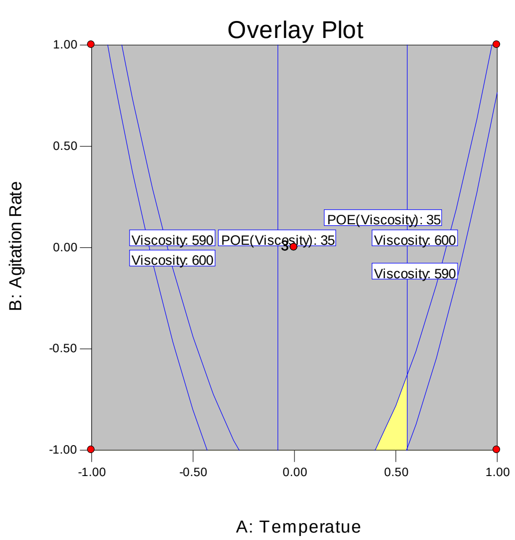
The corresponding contour and 3-D plots for this model are shown below followed by the POE contour and 3-D plots. Finally, the stacked contour plot is presented identifying a region with viscosity between 590 and 610 while minimizing the variability transmitted from the noise variable pressure. These conditions are in the region of factor *A* = 0.5 and factor *B* = -1.


Viscosity
B: Agitation Rate
A: Temperatue
660
640
620
600
580
560
540
1.000
0.500
0.000
-0.500
-1.000


Viscosity
675
650
625
600
575
550
525
A: Temperatue
B: Agitation Rate


POE(Viscosity)
B: Agitation Rate
A: Temperatue
43
42
41
40
39
38
37
36
35


POE(Viscosity)
44
42
40
38
36
34
A: Temperatue
B: Agitation Rate


Overlay Plot
B: Agitation Rate
A: Temperatue
Viscosity: 590
Viscosity: 600
POE(Viscosity): 35
POE(Viscosity): 35
Viscosity: 600
Viscosity: 590

12.9. A variation of Example 12.1. In example 12.1 (which utilized data from Example 6-2) we found that one of the process variables (B = pressure) was not important. Dropping this variable produced two replicates of a 2^3 design. The data are shown below.

C	D	$A(-)$	$A(+)$	$\bar{y}$	s^2
-	-	45, 48	71, 65	57.75	121.19
+	-	68, 80	60, 65	68.25	72.25
-	+	43, 45	100, 104	73.00	1124.67
+	+	75, 70	86, 96	81.75	134.92

Assume that C and D are controllable factors and that A is a noise factor.

(a) Fit a model to the mean response.

The following is the analysis of variance with all terms in the model:

Design Expert Output

Response: Mean
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	300.05	3	100.02		
A	*92.64*	*1*	*92.64*		
B	*206.64*	*1*	*206.64*		
AB	*0.77*	*1*	*0.77*		
Pure Error	0.000	0			
Cor Total	300.05	3			

Based on the above analysis, the AB interaction is removed from the model and used as error.

Design Expert Output

Response: Mean
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	299.28	2	149.64	195.45	0.0505	not significant
A	*92.64*	*1*	*92.64*	*121.00*	*0.0577*	
B	*206.64*	*1*	*206.64*	*269.90*	*0.0387*	
Residual	0.77	1	0.77			
Cor Total	300.05	3				

The Model F-value of 195.45 implies there is a 5.05% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	0.87	R-Squared	0.9974
Mean	70.19	Adj R-Squared	0.9923
C.V.	1.25	Pred R-Squared	0.9592
PRESS	12.25	Adeq Precision	31.672

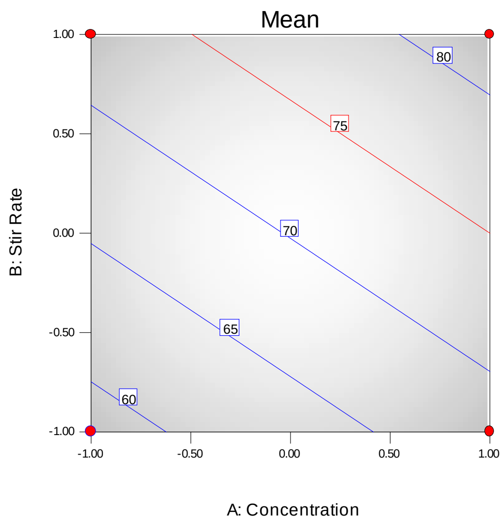
Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	70.19	1	0.44	64.63	75.75	
A-Concentration	4.81	1	0.44	-0.75	10.37	1.00
B-Stir Rate	7.19	1	0.44	1.63	12.75	1.00

Final Equation in Terms of Coded Factors:

Mean =
+70.19

```
                    +4.81    * A
                    +7.19    * B

Final Equation in Terms of Actual Factors:

                    Mean   =
                +70.18750
                 +4.81250    * Concentration
                 +7.18750    * Stir Rate
```

The following is a contour plot of the mean model:



(b) Fit a model to the $ln(s^2)$ response.

The following is the analysis of variance with all terms in the model:

Design Expert Output

Response: Variance Transform: Natural log Constant: 0
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	4.42	3	1.47		
A	*1.74*	*1*	*1.74*		
B	*2.03*	*1*	*2.03*		
AB	*0.64*	*1*	*0.64*		
Pure Error	0.000	0			
Cor Total	4.42	3			

Based on the above analysis, the AB interaction is removed from the model and applied to the residual error.

Design Expert Output

Response: Variance Transform: Natural log Constant: 0
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	3.77	2	1.89	2.94	0.3815	not significant
A	*1.74*	*1*	*1.74*	*2.71*	*0.3477*	
B	*2.03*	*1*	*2.03*	*3.17*	*0.3260*	
Residual	0.64	1	0.64			

Cor Total	4.42	3

The "Model F-value" of 2.94 implies the model is not significant relative to the noise. There is a 38.15 % chance that a "Model F-value" this large could occur due to noise.

Std. Dev.	0.80	R-Squared	0.8545
Mean	5.25	Adj R-Squared	0.5634
C.V.	15.26	Pred R-Squared	-1.3284
PRESS	10.28	Adeq Precision	3.954

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	5.25	1	0.40	0.16	10.34	
A-Concentration	-0.66	1	0.40	-5.75	4.43	1.00
B-Stir Rate	0.71	1	0.40	-4.38	5.81	1.00

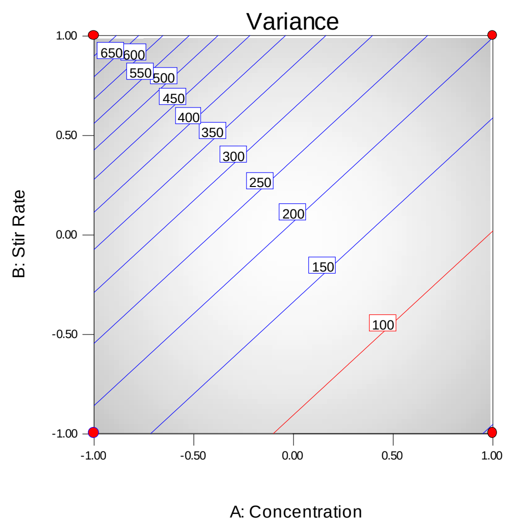
Final Equation in Terms of Coded Factors:

Ln(Variance) =
+5.25
-0.66 * A
+0.71 * B

Final Equation in Terms of Actual Factors:

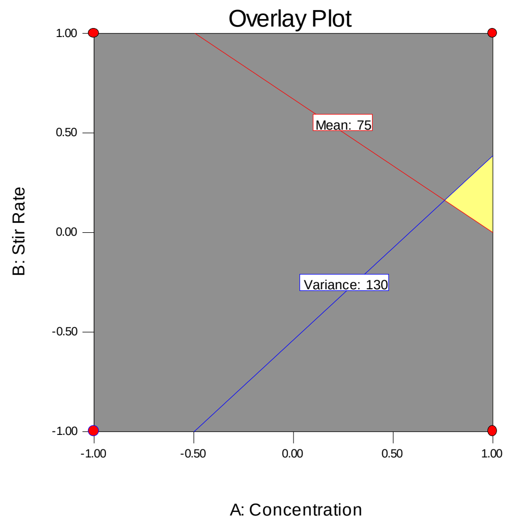
Ln(Variance) =
+5.25185
-0.65945 * Concentration
+0.71311 * Stir Rate

The following is a contour plot of the variance model in the untransformed form:



(c) Find operating conditions that result in the mean filtration rate response exceeding 75 with minimum variance.

The overlay plot shown below identifies the region required by the process:



(d) Compare your results with those from Example 12.1 which used the transmission of error approach. How similar are the two answers.

The results are very similar. Both require the Concentration to be held at the high level while the stirring rate is held near the middle.

12.10. In an article ("Let's All Beware the Latin Square," *Quality Engineering*, Vol. 1, 1989, pp. 453-465) J.S. Hunter illustrates some of the problems associated with 3^{k-p} fractional factorial designs. Factor *A* is the amount of ethanol added to a standard fuel and factor *B* represents the air/fuel ratio. The response variable is carbon monoxide (CO) emission in g/m^2. The design is shown below.

Design				Observations	
A	*B*	x_1	x_2	*y*	
0	0	-1	-1	66	62
1	0	0	-1	78	81
2	0	1	-1	90	94
0	1	-1	0	72	67
1	1	0	0	80	81
2	1	1	0	75	78
0	2	-1	1	68	66
1	2	0	1	66	69
2	2	1	1	60	58

Notice that we have used the notation system of 0, 1, and 2 to represent the low, medium, and high levels for the factors. We have also used a "geometric notation" of -1, 0, and 1. Each run in the design is replicated twice.

(a) Verify that the second-order model

$$\hat{y} = 78.5 + 4.5x_1 - 7.0x_2 - 4.5x_1^2 - 4.0x_2^2 - 9.0x_1x_2$$

is a reasonable model for this experiment. Sketch the CO concentration contours in the x_1, x_2 space.

In the computer output that follows, the "coded factors" model is in the -1, 0, +1 scale.

Design Expert Output

Response: CO Emis

ANOVA for Response Surface Quadratic Model

Analysis of variance table [Partial sum of squares]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	1624.00	5	324.80	50.95	< 0.0001	significant
A	*243.00*	*1*	*243.00*	*38.12*	*< 0.0001*	
B	*588.00*	*1*	*588.00*	*92.24*	*< 0.0001*	
A^2	*81.00*	*1*	*81.00*	*12.71*	*0.0039*	
B^2	*64.00*	*1*	*64.00*	*10.04*	*0.0081*	
AB	*648.00*	*1*	*648.00*	*101.65*	*< 0.0001*	
Residual	76.50	12	6.37			
Lack of Fit	*30.00*	*3*	*10.00*	*1.94*	*0.1944*	*not significant*
Pure Error	*46.50*	*9*	*5.17*			
Cor Total	1700.50	17				

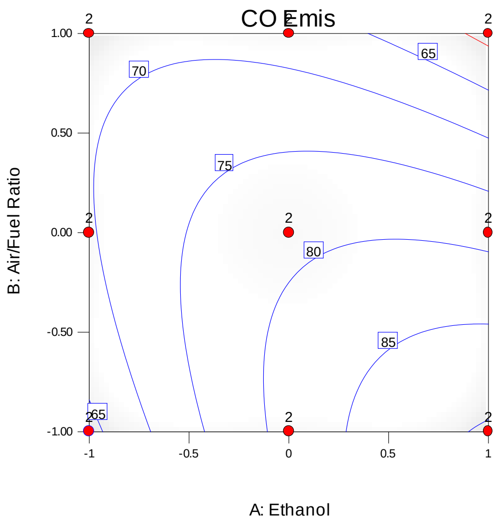
The Model F-value of 50.95 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.52	R-Squared	0.9550
Mean	72.83	Adj R-Squared	0.9363
C.V.	3.47	Pred R-Squared	0.9002
PRESS	169.71	Adeq Precision	21.952

Factor	**Coefficient Estimate**	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept	78.50	1	1.33	75.60	81.40	
A-Ethanol	4.50	1	0.73	2.91	6.09	1.00
B-Air/Fuel Ratio	-7.00	1	0.73	-8.59	-5.41	1.00
A^2	-4.50	1	1.26	-7.25	-1.75	1.00
B^2	-4.00	1	1.26	-6.75	-1.25	1.00
AB	-9.00	1	0.89	-10.94	-7.06	1.00

Final Equation in Terms of Coded Factors:

CO Emis =
+78.50
+4.50 * A
-7.00 * B
-4.50 * A^2
-4.00 * B^2
-9.00 * A * B



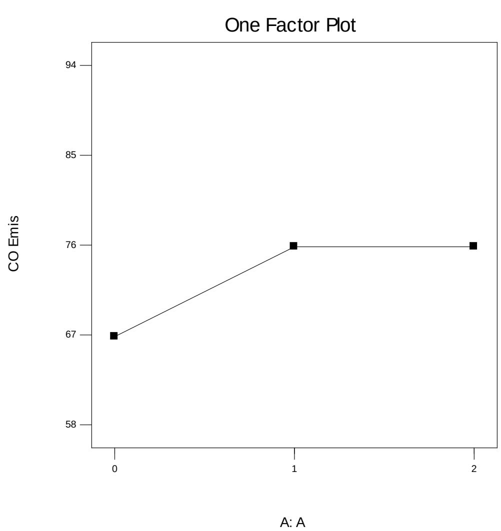
(b) Now suppose that instead of only two factors, we had used *four* factors in a 3^{4-2} fractional factorial design and obtained *exactly* the same data in part (a). The design would be as follows:

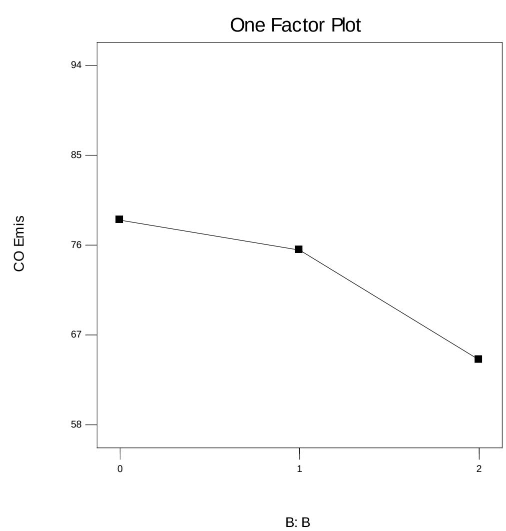
Design								Observations	
A	*B*	*C*	*D*	x_1	x_2	x_3	x_4	*y*	*y*
0	0	0	0	-1	-1	-1	-1	66	62
1	0	1	1	0	-1	0	0	78	81
2	0	2	2	+1	-1	+1	+1	90	94
0	1	2	1	-1	0	+1	0	72	67
1	1	0	2	0	0	-1	+1	80	81
2	1	1	0	+1	0	0	-1	75	78
0	2	1	2	-1	+1	0	+1	68	66
1	2	2	0	0	+1	+1	-1	66	69
2	2	0	1	+1	+1	-1	0	60	58

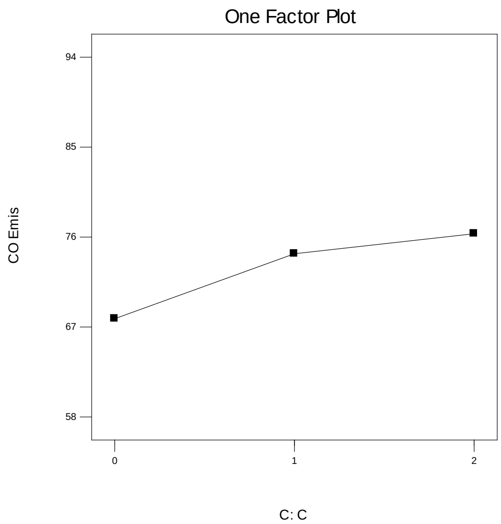
Calculate the marginal averages of the CO response at each level of four factors *A*, *B*, *C*, and *D*. Construct plots of these marginal averages and interpret the results. Do factors *C* and *D* appear to have strong effects? Do these factors *really* have any effect on CO emission? Why is their appearance effect strong?

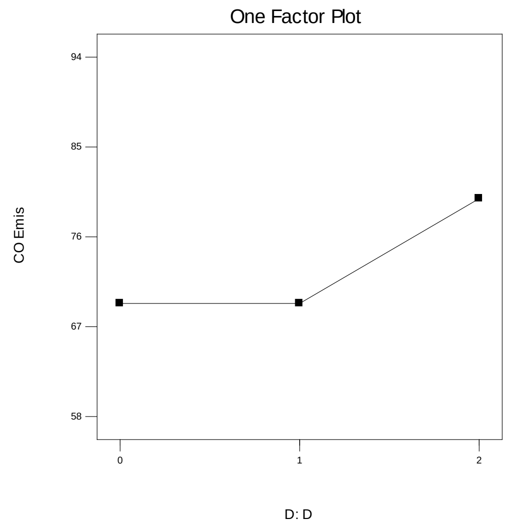
The marginal averages are shown below. Both Factors *C* and *D* appear to have an effect on CO emission. This is probably because both *C* and *D* are aliased with components of the interaction involving *A* and *B*, and there is a strong *AB* interaction.

Level	*A*	*B*	*C*	*D*
0	66.83	78.50	67.83	69.33
1	75.83	75.50	74.33	69.33
2	75.83	64.50	76.33	79.83









(c) The design in part (b) allows the model

$$y = \beta_0 + \sum_{i=1}^{4} \beta_i x_i + \sum_{i=1}^{4} \beta_{ii} x_i^2 + \varepsilon$$

to be fitted. Suppose that the *true* model is

$$y = \beta_0 + \sum_{i=1}^{4} \beta_i x_i + \sum_{i=1}^{4} \beta_{ii} x_i^2 + \sum\sum_{i<j} \beta_{ij} x_i x_j + \varepsilon$$

Show that if $\hat{\beta}_j$ represents the least squares estimates of the coefficients in the fitted model, then

$$\begin{aligned}
E(\hat{\beta}_0) &= \beta_0 - \beta_{13} - \beta_{14} - \beta_{34} \\
E(\hat{\beta}_1) &= \beta_1 - (\beta_{23} + \beta_{24})/2 \\
E(\hat{\beta}_2) &= \beta_2 - (\beta_{13} + \beta_{14} + \beta_{34})/2 \\
E(\hat{\beta}_3) &= \beta_3 - (\beta_{12} + \beta_{24})/2 \\
E(\hat{\beta}_4) &= \beta_4 - (\beta_{12} + \beta_{23})/2 \\
E(\hat{\beta}_{11}) &= \beta_{11} - (\beta_{23} - \beta_{24})/2 \\
E(\hat{\beta}_{22}) &= \beta_{22} + (\beta_{13} + \beta_{14} + \beta_{34})/2 \\
E(\hat{\beta}_{33}) &= \beta_{33} - (\beta_{24} - \beta_{12})/2 + \beta_{14} \\
E(\hat{\beta}_{44}) &= \beta_{44} - (\beta_{12} - \beta_{23})/2 + \beta_{13}
\end{aligned}$$

Does this help explain the strong effects for factors *C* and *D* observed graphically in part (b)?

$$\text{Let } \mathbf{X}_1 = \begin{array}{c} \begin{matrix} \beta_0 & \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_{11} & \beta_{22} & \beta_{33} & \beta_{44} \end{matrix} \\ \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 & 1 & 0 & 0 & 1 \\ 1 & -1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & -1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

$$\text{and } \mathbf{X}_2 = \begin{array}{c} \begin{matrix} \beta_{12} & \beta_{13} & \beta_{14} & \beta_{23} & \beta_{24} & \beta_{34} \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & -1 & -1 & 1 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ 1 & -1 & 0 & -1 & 0 & 0 \end{bmatrix} \end{array}$$

$$\text{Then, } \mathbf{A} = \left(\mathbf{X}_1'\mathbf{X}_1\right)^{-1}\mathbf{X}_1'\mathbf{X}_2 = \begin{bmatrix} 0 & -1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1/2 & -1/2 & 0 \\ 0 & -1/2 & -1/2 & 0 & 0 & -1/2 \\ -1/2 & 0 & 0 & 0 & -1/2 & 0 \\ -1/2 & 0 & 0 & -1/2 & 0 & 0 \\ 0 & 0 & 0 & -1/2 & 1/2 & 1 \\ 0 & 1/2 & 1/2 & 0 & 0 & 1/2 \\ 1/2 & 0 & 1 & 0 & -1/2 & 0 \\ -1/2 & 1 & 0 & 1/2 & 0 & 0 \end{bmatrix}$$

$$E\begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\\ \hat{\beta}_4\\ \hat{\beta}_{11}\\ \hat{\beta}_{22}\\ \hat{\beta}_{33}\\ \hat{\beta}_{44}\end{bmatrix}=\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_{11}\\ \beta_{22}\\ \beta_{33}\\ \beta_{44}\end{bmatrix}+\begin{bmatrix}0 & -1 & -1 & 0 & 0 & -1\\ 0 & 0 & 0 & -1/2 & -1/2 & 0\\ 0 & -1/2 & -1/2 & 0 & 0 & -1/2\\ -1/2 & 0 & 0 & 0 & -1/2 & 0\\ -1/2 & 0 & 0 & -1/2 & 0 & 0\\ 0 & 0 & 0 & -1/2 & 1/2 & 1\\ 0 & 1/2 & 1/2 & 0 & 0 & 1/2\\ 1/2 & 0 & 1 & 0 & -1/2 & 0\\ -1/2 & 1 & 0 & 1/2 & 0 & 0\end{bmatrix}\begin{bmatrix}\beta_{12}\\ \beta_{13}\\ \beta_{14}\\ \beta_{23}\\ \beta_{24}\\ \beta_{34}\end{bmatrix}=\begin{bmatrix}\beta_0-\beta_{13}-\beta_{14}-\beta_{34}\\ \beta_1-1/2\beta_{23}-1/2\beta_{24}\\ \beta_2-1/2\beta_{13}-1/2\beta_{14}-1/2\beta_{34}\\ \beta_3-1/2\beta_{12}-1/2\beta_{24}\\ \beta_4-1/2\beta_{12}-1/2\beta_{23}\\ \beta_{11}-1/2\beta_{23}+1/2\beta_{24}+\beta_{34}\\ \beta_{22}+1/2\beta_{13}+1/2\beta_{14}+1/2\beta_{34}\\ \beta_{33}+1/2\beta_{12}+\beta_{14}-1/2\beta_{24}\\ \beta_{44}-1/2\beta_{12}+\beta_{13}+1/2\beta_{23}\end{bmatrix}$$

12.11. An experiment has been run in a process that applies a coating material to a wafer. Each run in the experiment produced a wafer, and the coating thickness was measured several times at different locations on the wafer. Then the mean y_1, and standard deviation y_2 of the thickness measurement was obtained. The data [adapted from Box and Draper (1987)] are shown in the table below.

Run	Speed	Pressure	Distance	Mean (y_1)	Std Dev (y_2)
1	-1.000	-1.000	-1.000	24.0	12.5
2	0.000	-1.000	-1.000	120.3	8.4
3	1.000	-1.000	-1.000	213.7	42.8
4	-1.000	0.000	-1.000	86.0	3.5
5	0.000	0.000	-1.000	136.6	80.4
6	1.000	0.000	-1.000	340.7	16.2
7	-1.000	1.000	-1.000	112.3	27.6
8	0.000	1.000	-1.000	256.3	4.6
9	1.000	1.000	-1.000	271.7	23.6
10	-1.000	-1.000	0.000	81.0	0.0
11	0.000	-1.000	0.000	101.7	17.7
12	1.000	-1.000	0.000	357.0	32.9
13	-1.000	0.000	0.000	171.3	15.0
14	0.000	0.000	0.000	372.0	0.0
15	1.000	0.000	0.000	501.7	92.5
16	-1.000	1.000	0.000	264.0	63.5
17	0.000	1.000	0.000	427.0	88.6
18	1.000	1.000	0.000	730.7	21.1
19	-1.000	-1.000	1.000	220.7	133.8
20	0.000	-1.000	1.000	239.7	23.5
21	1.000	-1.000	1.000	422.0	18.5
22	-1.000	0.000	1.000	199.0	29.4
23	0.000	0.000	1.000	485.3	44.7
24	1.000	0.000	1.000	673.7	158.2
25	-1.000	1.000	1.000	176.7	55.5
26	0.000	1.000	1.000	501.0	138.9
27	1.000	1.000	1.000	1010.0	142.4

(a) What type of design did the experimenters use? Is this a good choice of design for fitting a quadratic model?

The design is a 3^3. A better choice would be a 2^3 central composite design. The CCD gives more information over the design region with fewer points.

(b) Build models of both responses.

The model for the mean is developed as follows:

Design Expert Output

Response: Mean

ANOVA for Response Surface Reduced Cubic Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	1.289E+006	7	1.841E+005	60.45	< 0.0001	significant
A	*5.640E+005*	*1*	*5.640E+005*	*185.16*	*< 0.0001*	
B	*2.155E+005*	*1*	*2.155E+005*	*70.75*	*< 0.0001*	
C	*3.111E+005*	*1*	*3.111E+005*	*102.14*	*< 0.0001*	
AB	*52324.81*	*1*	*52324.81*	*17.18*	*0.0006*	
AC	*68327.52*	*1*	*68327.52*	*22.43*	*0.0001*	
BC	*22794.08*	*1*	*22794.08*	*7.48*	*0.0131*	
ABC	*54830.16*	*1*	*54830.16*	*18.00*	*0.0004*	
Residual	57874.57	19	3046.03			
Cor Total	1.347E+006	26				

The Model F-value of 60.45 implies the model is significant. There is only
a 0.01% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	55.19	R-Squared	0.9570
Mean	314.67	Adj R-Squared	0.9412
C.V.	17.54	Pred R-Squared	0.9056
PRESS	1.271E+005	Adeq Precision	33.333

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	314.67	1	10.62	292.44	336.90	
A-Speed	177.01	1	13.01	149.78	204.24	1.00
B-Pressure	109.42	1	13.01	82.19	136.65	1.00
C-Distance	131.47	1	13.01	104.24	158.70	1.00
AB	66.03	1	15.93	32.69	99.38	1.00
AC	75.46	1	15.93	42.11	108.80	1.00
BC	43.58	1	15.93	10.24	76.93	1.00
ABC	82.79	1	19.51	41.95	123.63	1.00

Final Equation in Terms of Coded Factors:

Mean =
+314.67
+177.01 * A
+109.42 * B
+131.47 * C
+66.03 * A * B
+75.46 * A * C
+43.58 * B * C
+82.79 * A * B * C

Final Equation in Terms of Actual Factors:

Mean =
+314.67037
+177.01111 * Speed
+109.42222 * Pressure
+131.47222 * Distance
+66.03333 * Speed * Pressure
+75.45833 * Speed * Distance
+43.58333 * Pressure * Distance
+82.78750 * Speed * Pressure * Distance

The model for the Std. Dev. response is as follows. A square root transformation was applied to correct problems with the normality assumption.

Design Expert Output

Response: Std. Dev. Transform: Square root Constant: 0

ANOVA for Response Surface Linear Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	116.75	3	38.92	4.34	0.0145	significant
A	*16.52*	*1*	*16.52*	*1.84*	*0.1878*	
B	*26.32*	*1*	*26.32*	*2.94*	*0.1001*	
C	*73.92*	*1*	*73.92*	*8.25*	*0.0086*	
Residual	206.17	23	8.96			
Cor Total	322.92	26				

The Model F-value of 4.34 implies the model is significant. There is only
a 1.45% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	2.99	R-Squared	0.3616
Mean	6.00	Adj R-Squared	0.2783
C.V.	49.88	Pred R-Squared	0.1359
PRESS	279.05	Adeq Precision	7.278

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	6.00	1	0.58	4.81	7.19	
A-Speed	0.96	1	0.71	-0.50	2.42	1.00
B-Pressure	1.21	1	0.71	-0.25	2.67	1.00
C-Distance	2.03	1	0.71	0.57	3.49	1.00

Final Equation in Terms of Coded Factors:

Sqrt(Std. Dev.) =
+6.00
+0.96 * A
+1.21 * B
+2.03 * C

Final Equation in Terms of Actual Factors:

Sqrt(Std. Dev.) =
+6.00273
+0.95796 * Speed
+1.20916 * Pressure
+2.02643 * Distance

Because Factor *A* is insignificant, it is removed from the model. The reduced linear model analysis is shown below:

Design Expert Output

Response: Std. Dev. Transform: Square root Constant: 0

ANOVA for Response Surface Reduced Linear Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F	
Model	100.23	2	50.12	5.40	0.0116	significant
B	*26.32*	*1*	*26.32*	*2.84*	*0.1051*	
C	*73.92*	*1*	*73.92*	*7.97*	*0.0094*	
Residual	222.68	24	9.28			
Cor Total	322.92	26				

The Model F-value of 5.40 implies the model is significant. There is only
a 1.16% chance that a "Model F-Value" this large could occur due to noise.

Std. Dev.	3.05	R-Squared	0.3104
Mean	6.00	Adj R-Squared	0.2529
C.V.	50.74	Pred R-Squared	0.1476
PRESS	275.24	Adeq Precision	6.373

Factor	Coefficient Estimate	DF	Standard Error	95% CI Low	95% CI High	VIF
Intercept	6.00	1	0.59	4.79	7.21	
B-Pressure	1.21	1	0.72	-0.27	2.69	1.00
C-Distance	2.03	1	0.72	0.54	3.51	1.00

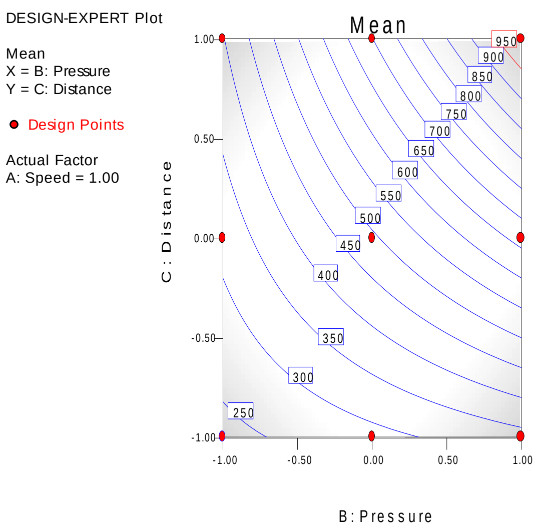
Final Equation in Terms of Coded Factors:

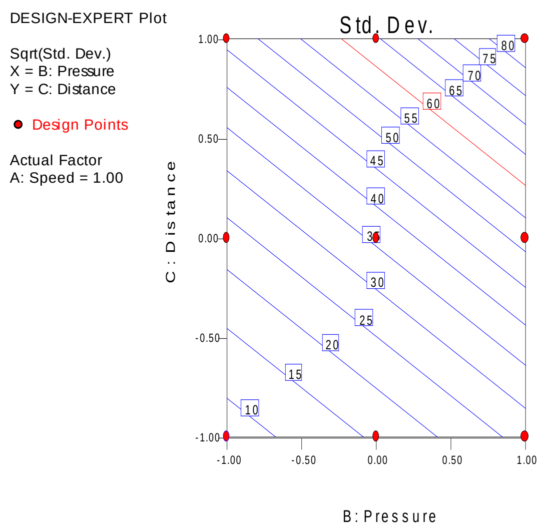
Sqrt(Std. Dev.) =
+6.00
+1.21 * B
+2.03 * C

Final Equation in Terms of Actual Factors:

Sqrt(Std. Dev.) =
+6.00273
+1.20916 * Pressure
+2.02643 * Distance

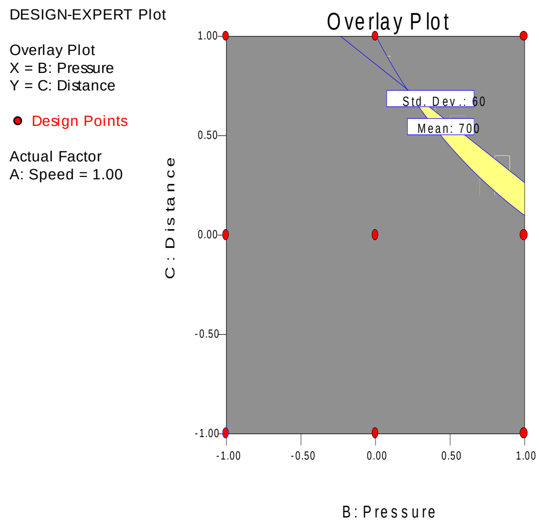
The following contour plots graphically represent the two models.





(c) Find a set of optimum conditions that result in the mean as large as possible with the standard deviation less than 60.

The overlay plot identifies a region that meets the criteria of the mean as large as possible with the standard deviation less than 60. The optimum conditions in coded terms are approximately Speed = 1.0, Pressure = 0.75 and Distance = 0.25.



12.12. Suppose that there are four controllable variables and two noise variables. It is necessary to estimate the main effects and two-factor interactions of all of the controllable variables, the main effects of the noise variables, and the two-factor interactions between all controllable and noise factors. If all factors are at two levels, what is the minimum number of runs that can be used to estimate all of the model parameters using a combined array design? Use a D-optimal algorithm to find a design.

Twenty-one runs are required for the model, with five additional runs for lack of fit, and five as replicates for a total of 31 runs as follows. It should be noted that *Design Expert's D*-optimal algorithm might not create the same design if repeated.

Std	*A*	*B*	*C*	*D*	*E*	*F*
1	+1	+1	-1	+1	+1	+1
2	-1	+1	-1	+1	-1	-1
3	+1	-1	-1	+1	-1	-1
4	+1	+1	-1	-1	-1	+1
5	-1	+1	-1	-1	+1	+1
6	-1	+1	+1	+1	+1	+1
7	+1	+1	-1	-1	+1	-1
8	-1	-1	+1	+1	-1	-1
9	-1	+1	+1	-1	+1	-1
10	-1	+1	+1	-1	-1	+1
11	+1	-1	+1	+1	+1	+1
12	+1	+1	+1	+1	-1	+1
13	+1	-1	-1	-1	+1	+1
14	+1	+1	+1	-1	+1	+1
15	-1	-1	-1	-1	-1	-1
16	+1	+1	+1	+1	+1	-1
17	-1	-1	-1	+1	-1	+1
18	-1	-1	-1	+1	+1	-1
19	-1	-1	+1	-1	+1	+1
20	+1	-1	+1	-1	+1	-1
21	+1	-1	+1	-1	-1	+1
22	+1	+1	+1	-1	-1	-1
23	+1	-1	-1	-1	-1	-1
24	-1	+1	-1	-1	-1	-1
25	+1	+1	-1	-1	-1	-1
26	-1	-1	+1	-1	-1	-1
27	+1	+1	+1	+1	-1	+1
28	-1	-1	-1	+1	-1	+1
29	+1	+1	+1	+1	+1	-1
30	-1	-1	-1	+1	+1	-1
31	-1	+1	-1	-1	+1	+1

12.13. Suppose that there are four controllable variables and two noise variables. It is necessary to fit a complete quadratic model in the controllable variables, the main effects of the noise variables, and the two-factor interactions between all controllable and noise factors. Set up a combined array design for this by modifying a central composite design.

The following design is a half fraction central composite design with the axial points removed from the noise factors. The total number of runs is forty-eight which includes eight center points.

Std	*A*	*B*	*C*	*D*	*E*	*F*
1	-1	-1	-1	-1	-1	-1
2	+1	-1	-1	-1	-1	+1
3	-1	+1	-1	-1	-1	+1
4	+1	+1	-1	-1	-1	-1
5	-1	-1	+1	-1	-1	+1
6	+1	-1	+1	-1	-1	-1
7	-1	+1	+1	-1	-1	-1
8	+1	+1	+1	-1	-1	+1
9	-1	-1	-1	+1	-1	+1
10	+1	-1	-1	+1	-1	-1
11	-1	+1	-1	+1	-1	-1
12	+1	+1	-1	+1	-1	+1
13	-1	-1	+1	+1	-1	-1
14	+1	-1	+1	+1	-1	+1
15	-1	+1	+1	+1	-1	+1
16	+1	+1	+1	+1	-1	-1
17	-1	-1	-1	-1	+1	+1
18	+1	-1	-1	-1	+1	-1
19	-1	+1	-1	-1	+1	-1
20	+1	+1	-1	-1	+1	+1
21	-1	-1	+1	-1	+1	-1
22	+1	-1	+1	-1	+1	+1
23	-1	+1	+1	-1	+1	+1
24	+1	+1	+1	-1	+1	-1
25	-1	-1	-1	+1	+1	-1
26	+1	-1	-1	+1	+1	+1
27	-1	+1	-1	+1	+1	+1
28	+1	+1	-1	+1	+1	-1
29	-1	-1	+1	+1	+1	+1
30	+1	-1	+1	+1	+1	-1
31	-1	+1	+1	+1	+1	-1
32	+1	+1	+1	+1	+1	+1
33	-2.378	0	0	0	0	0
34	+2.378	0	0	0	0	0
35	0	-2.378	0	0	0	0
36	0	+2.378	0	0	0	0
37	0	0	-2.378	0	0	0
38	0	0	+2.378	0	0	0
39	0	0	0	-2.378	0	0
40	0	0	0	+2.378	0	0
41	0	0	0	0	0	0
42	0	0	0	0	0	0
43	0	0	0	0	0	0
44	0	0	0	0	0	0
45	0	0	0	0	0	0
46	0	0	0	0	0	0
47	0	0	0	0	0	0
48	0	0	0	0	0	0

12.14. Reconsider the situation in Problem 12.13. Could a modified small composite design be used for this problem? Are there any disadvantages associated with the use of the small composite design?

The axial points for the noise factors were removed in following small central composite design. Five center points are included. The small central composite design does have aliasing with noise factor *E* aliased with the *AD* interaction and noise factor *F* aliased with the *BC* interaction. These aliases are corrected by leaving the axial points for the noise factors in the design.

Std	*A*	*B*	*C*	*D*	*E*	*F*
1	+1	+1	+1	+1	-1	-1
2	+1	+1	+1	-1	+1	-1
3	+1	+1	-1	+1	-1	+1
4	+1	-1	+1	-1	+1	+1
5	-1	+1	-1	+1	+1	+1
6	+1	-1	+1	+1	-1	+1
7	-1	+1	+1	-1	-1	-1
8	+1	+1	-1	-1	+1	+1
9	+1	-1	-1	+1	-1	-1
10	-1	-1	+1	-1	-1	+1
11	-1	+1	-1	-1	-1	+1
12	+1	-1	-1	-1	+1	-1
13	-1	-1	-1	+1	+1	-1
14	-1	-1	+1	+1	+1	+1
15	-1	+1	+1	+1	+1	-1
16	-1	-1	-1	-1	-1	-1
17	-2	0	0	0	0	0
18	+2	0	0	0	0	0
19	0	-2	0	0	0	0
20	0	+2	0	0	0	0
21	0	0	-2	0	0	0
22	0	0	+2	0	0	0
23	0	0	0	-2	0	0
24	0	0	0	+2	0	0
25	0	0	0	0	0	0
26	0	0	0	0	0	0
27	0	0	0	0	0	0
28	0	0	0	0	0	0
29	0	0	0	0	0	0

12.15. Reconsider the situation in Problem 12.13. What is the minimum number of runs that can be used to estimate all of the model parameters using a combined array design? Use a D-optimal algorithm to find a reasonable design for this problem.

A minimum of 25 runs is required. The following design is a 36 run D-optimal with six additional runs included for lack of fit and five as replicates. It should be noted that Design Expert's *D*-optimal algorithm might not create the same design if repeated.

Std	*A*	*B*	*C*	*D*	*E*	*F*
1	+1	+1	+1	-1	-1	-1
2	-1	+1	-1	-1	+1	+1
3	-1	+1	+1	+1	-1	+1
4	+1	+1	-1	+1	-1	-1
5	-1	-1	+1	+1	-1	-1
6	-1	+1	-1	-1	-1	-1
7	+1	-1	-1	+1	+1	-1
8	+1	-1	+1	-1	+1	-1
9	+1	+1	-1	+1	+1	+1
10	+1	-1	-1	-1	-1	-1
11	+1	-1	+1	+1	-1	+1
12	-1	+1	-1	+1	+1	-1
13	+1	+1	+1	-1	+1	+1
14	+1	+1	-1	-1	-1	+1
15	+1	+1	+1	+1	+1	-1
16	-1	-1	-1	+1	-1	+1
17	0	-1	-1	-1	+1	-1
18	0	-1	+1	-1	-1	+1
19	0	+1	0	0	0	0
20	0	0	0	-1	0	0
21	0	0	+1	0	0	0

22	-1	+1	+1	-1	+1	-1
23	-1	-1	+1	0	+1	+1
24	+1	+1	-1	-1	+1	-1
25	0	-1	+1	+1	+1	+1
26	+1	-1	-1	-1	+1	+1
27	-1	-1	-1	0	-1	-1
28	+1	-1	0	+1	-1	-1
29	-1	-1	0	+1	+1	-1
30	+1	-1	-1	0	-1	+1
31	-1	0	-1	+1	+1	+1
32	+1	+1	+1	+1	+1	-1
33	+1	-1	+1	-1	+1	-1
34	-1	+1	+1	+1	-1	+1
35	+1	+1	+1	-1	-1	-1
36	+1	+1	-1	+1	+1	+1

12.16. Rework Problem 12.15 using the *I*-criterion to construct the design. Compare this design to the *D*-optimal design in Problem 12.15. Which design would you prefer?

The JMP Output shown below identifies an *I*-optimal design with 25 runs. It should be noted that JMP's *I*-optimal design algorithm might not create the same design if repeated. We would prefer the *I*-optimal over the *D*-optimal because a response surface model is of interest in this problem.

JMP Output

Custom Design
Factors
Add N Factors
1

X1 Continuous
X2 Continuous
X3 Continuous
X4 Continuous
Noise1 Continuous
Noise2 Continuous

Design

Run	X1	X2	X3	X4	Noise1	Noise2
1	1	-1	0	-1	-1	1
2	-1	1	-1	1	0	1
3	0	0	1	1	1	-1
4	1	-1	-1	-1	-1	-1
5	1	1	1	-1	-1	-1
6	1	1	-1	0	-1	1
7	-1	1	1	1	-1	-1
8	1	-1	1	1	0	1
9	-1	1	0	-1	1	1
10	-1	-1	-1	0	-1	1
11	-1	-1	0	-1	0	-1
12	1	0	0	0	1	-1
13	0	0	0	1	-1	1
14	-1	-1	0	1	1	1
15	0	-1	1	0	-1	-1
16	-1	1	1	-1	1	-1
17	1	1	-1	1	1	-1
18	-1	0	1	-1	-1	0
19	0	-1	1	-1	1	1
20	0	1	0	-1	-1	-1
21	1	-1	-1	1	1	1
22	1	1	-1	-1	1	1
23	0	0	-1	0	1	-1
24	-1	-1	-1	1	-1	-1
25	0	1	1	0	1	1

12.17. Rework Problem 12.12 using the *I*-criterion. Compare this to the D-optimal design in Problem 12.12. Which design would you prefer?

The JMP Output shown below identifies an *I*-optimal design with 21 runs. It should be noted that JMP's *I*-optimal design algorithm might not create the same design if repeated. Because the problem requires the experiment be run in at only two levels for each variable, and I-optimal algorithm in JMP generates center points for continuous variables, categorical variables were used. We would prefer the *D*-optimal over the *I*-optimal assuming that the interest is to identify the important factors more so than to fit a response surface model. Also, the *D*-optimal algorithm in Design Expert will generate a design for this model with continuous variables and still maintain the requirement for running the experiment at only two levels for each variable.

JMP Output

Custom Design
Factors
Add N Factors
1

X1 Continuous
X2 Continuous
X3 Continuous
X4 Continuous
Noise1 Continuous
Noise2 Continuous

Design

Run	X1	X2	X3	X4	Noise1	Noise2
1	1	1	1	1	-1	-1
2	-1	1	-1	-1	1	-1
3	-1	1	-1	1	1	-1
4	-1	1	-1	-1	-1	-1
5	-1	-1	1	1	-1	-1
6	1	-1	-1	1	-1	-1
7	1	1	1	-1	1	-1
8	-1	1	1	1	-1	-1
9	-1	-1	1	-1	-1	1
10	-1	-1	-1	1	-1	-1
11	1	-1	-1	1	-1	-1
12	1	-1	1	-1	-1	-1
13	1	1	-1	-1	1	1
14	1	1	1	1	-1	1
15	-1	1	1	-1	1	1
16	-1	-1	1	-1	-1	1
17	-1	-1	1	1	-1	1
18	-1	1	1	1	1	-1
19	-1	-1	-1	-1	1	1
20	1	1	-1	-1	-1	1
21	-1	-1	-1	1	-1	1
22	1	1	-1	-1	1	1

12.18. An experiment was run in a wave soldering process. There are five controllable variables and three noise variables. The response variable is the number of solder defects per million opportunities. The experimental design employed was the crossed array shown below.

					Outer Array				
					F	-1	1	1	-1
Inner Array					*G*	-1	1	-1	1
A	*B*	*C*	*D*	*E*	*H*	-1	-1	1	1
1	1	1	-1	-1		194	197	193	275
1	1	-1	1	1		136	136	132	136
1	-1	1	-1	1		185	261	264	264
1	-1	-1	1	-1		47	125	127	42
-1	1	1	1	-1		295	216	204	293
-1	1	-1	-1	1		234	159	231	157
-1	-1	1	1	1		328	326	247	322
-1	-1	-1	-1	-1		186	187	105	104

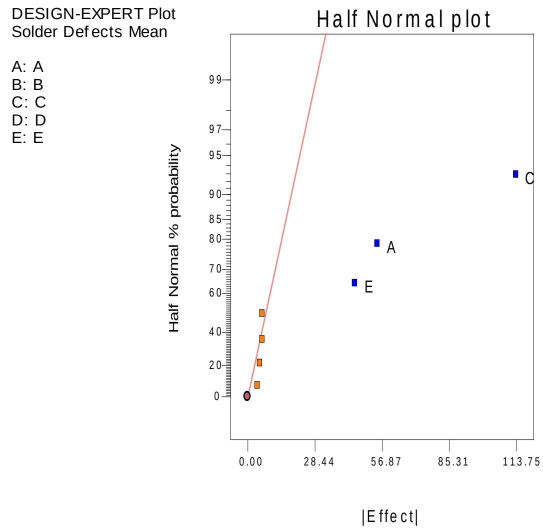
(a) What types of designs were used for the inner and outer arrays?

The inner array is a 2^{5-2} fractional factorial design with a defining relation of $I = -ACD = -BCE = ABDE$. The outer array is a 2^{3-1} fractional factorial design with a defining relation of $I = -FGH$.

(b) Develop models for the mean and variance of solder defects. What set of operating conditions would you recommend?

A	*B*	*C*	*D*	*E*	$\bar{y}$	s^2
1	1	1	-1	-1	214.75	1616.25
1	1	-1	1	1	135.00	4.00
1	-1	1	-1	1	243.50	1523.00
1	-1	-1	1	-1	85.25	2218.92
-1	1	1	1	-1	252.00	2376.67
-1	1	-1	-1	1	195.25	1852.25
-1	-1	1	1	1	305.75	1540.25
-1	-1	-1	-1	-1	145.50	2241.67

The following analysis identifies factors A, C, and E as being significant for the solder defects mean model.



Design Expert Output

Response: Solder Defects Mean
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	36068.63	3	12022.88	194.31	< 0.0001	significant
A	*6050.00*	*1*	*6050.00*	*97.78*	*0.0006*	
C	*25878.13*	*1*	*25878.13*	*418.23*	*< 0.0001*	
E	*4140.50*	*1*	*4140.50*	*66.92*	*0.0012*	
Residual	247.50	4	61.88			
Cor Total	36316.13	7				

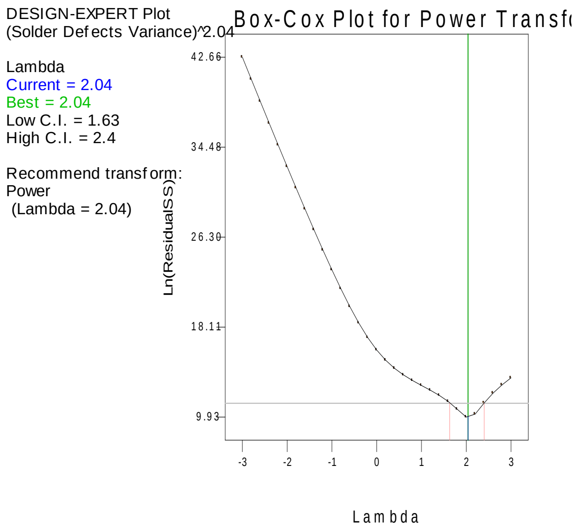
The "Model F-value" of 194.31implies the model is not significant relative to the noise. There is a 0.01 % chance that a "Model F-value" this large could occur due to noise.

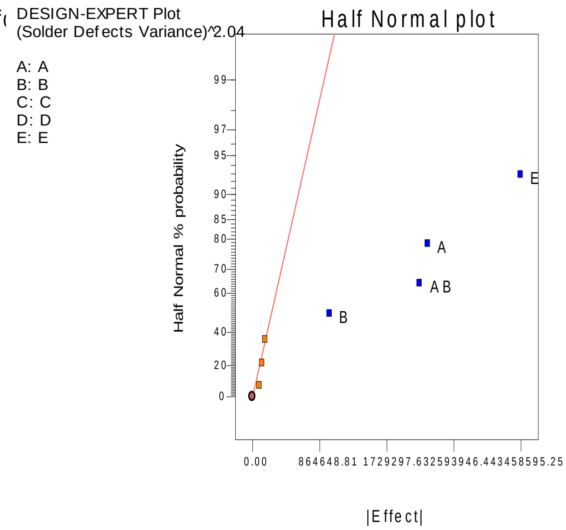
Std. Dev.	7.87	R-Squared	0.9932
Mean	197.13	Adj R-Squared	0.9881
C.V.	3.99	Pred R-Squared	0.9727
PRESS	990.00	Adeq Precision	38.519

Final Equation in Terms of Coded Factors:

Solder Defects Mean =
+197.13
-27.50 * A
+56.88 * C
+22.75 * E

Although the natural log transformation is often utilized for variance response, a power transformation actually performed better for this problem per the Box-Cox plot below. The analysis for the solder defect variance follows.





Design Expert Output

Response: Solder Defects Variance Transform: Power Lambda: 2.04 Constant: 0

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	4.542E+013	4	1.136E+013	325.30	0.0003	significant
A	*1.023E+013*	*1*	*1.023E+013*	*293.08*	*0.0004*	
B	*1.979E+012*	*1*	*1.979E+012*	*56.70*	*0.0049*	
E	*2.392E+013*	*1*	*2.392E+013*	*685.33*	*0.0001*	
AB	*9.289E+012*	*1*	*9.289E+012*	*266.11*	*0.0005*	
Residual	1.047E+011	3	3.491E+010			
Cor Total	4.553E+013	7				

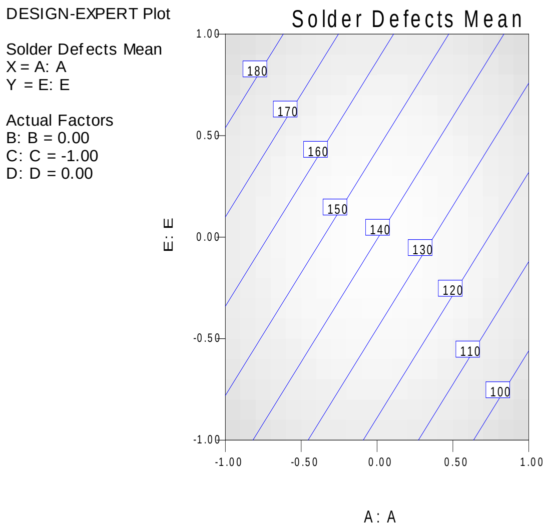
The "Model F-value" of 325.30 implies the model is not significant relative to the noise. There is a 0.03 % chance that a "Model F-value" this large could occur due to noise.

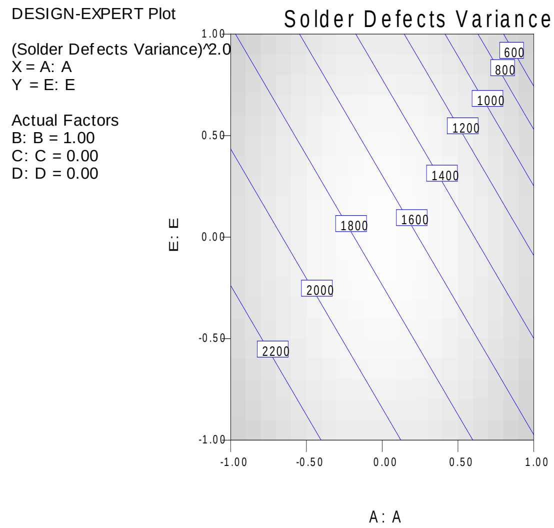
Std. Dev.	1.868E+005	R-Squared	0.9977
Mean	4.461E+006	Adj R-Squared	0.9946
C.V.	4.19	Pred R-Squared	0.9836
PRESS	7.447E+011	Adeq Precision	53.318

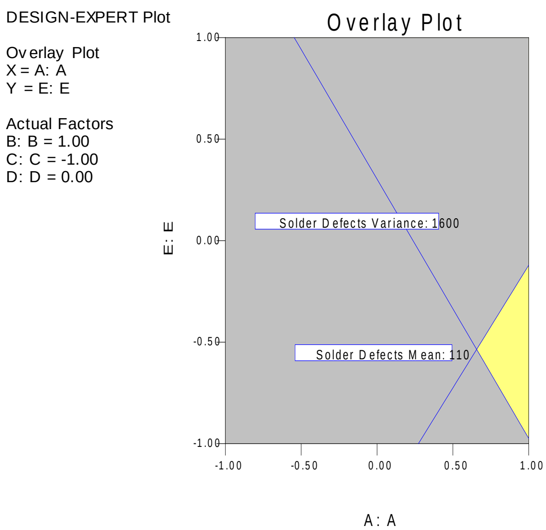
Final Equation in Terms of Coded Factors:

(Solder Defects Variance)2.04 =
+4.461E+006
-1.131E+006 * A
-4.974E+005 * B
-1.729E+006 * E
-1.078E+006 * A * B

The contour plots of the mean and variance models are shown below along with the overlay plot. Assuming that we wish to minimize both solder defects mean and variance, a solution is shown in the overlay plot with factors A = +1, B = +1, C = -1, D = 0, and E near -1.







12.19. Reconsider the wave soldering experiment in Problem 12.18. Find a combined array design for this experiment that requires fewer runs.

The following experiment is a 2^{8-4}, resolution IV design with the defining relation $I = BCDE = ACDF = ABCG = ABDH$. Only 16 runs are required.

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*
-1	-1	-1	-1	-1	-1	-1	-1
+1	-1	-1	-1	-1	+1	+1	+1
-1	+1	-1	-1	+1	-1	+1	+1
+1	+1	-1	-1	+1	+1	-1	-1
-1	-1	+1	-1	+1	+1	+1	-1
+1	-1	+1	-1	+1	-1	-1	+1
-1	+1	+1	-1	-1	+1	-1	+1
+1	+1	+1	-1	-1	-1	+1	-1
-1	-1	-1	+1	+1	+1	-1	+1

+1	-1	-1	+1	+1	-1	+1	-1
-1	+1	-1	+1	-1	+1	+1	-1
+1	+1	-1	+1	-1	-1	-1	+1
-1	-1	+1	+1	-1	-1	+1	+1
+1	-1	+1	+1	-1	+1	-1	-1
-1	+1	+1	+1	+1	-1	-1	-1
+1	+1	+1	+1	+1	+1	+1	+1

12.20. Reconsider the wave soldering experiment in Problem 12.18. Suppose that it was necessary to fit a complete quadratic model in the controllable variables, all main effects of the noise variables, and all controllable variable-noise variable interactions. What design would you recommend?

The following experiment is a small central composite design with five center points; the axial points for the noise factors have been removed. A total of 45 runs are required.

A	*B*	*C*	*D*	*E*	*F*	*G*	*H*
+1	+1	+1	-1	-1	+1	+1	+1
-1	+1	+1	-1	+1	+1	+1	-1
+1	+1	-1	-1	+1	+1	-1	-1
+1	-1	-1	+1	+1	-1	+1	-1
-1	-1	+1	+1	+1	+1	-1	-1
-1	+1	+1	+1	-1	+1	-1	-1
-1	+1	+1	+1	+1	+1	-1	-1
+1	+1	+1	+1	+1	-1	+1	-1
+1	+1	-1	+1	-1	+1	-1	+1
+1	-1	+1	+1	-1	-1	+1	-1
+1	+1	+1	-1	+1	-1	-1	+1
-1	+1	+1	-1	-1	-1	+1	+1
+1	-1	-1	-1	-1	+1	-1	-1
+1	-1	+1	-1	-1	-1	-1	+1
-1	+1	-1	-1	+1	+1	+1	-1
+1	-1	-1	+1	+1	+1	-1	+1
-1	-1	-1	-1	-1	+1	+1	-1
-1	-1	+1	+1	+1	-1	-1	+1
-1	+1	-1	-1	+1	-1	+1	+1
-1	-1	+1	+1	-1	+1	+1	+1
+1	+1	-1	+1	-1	-1	+1	-1
-1	-1	+1	-1	+1	-1	-1	-1
+1	+1	+1	-1	-1	+1	-1	-1
+1	-1	-1	-1	+1	+1	+1	+1
+1	-1	+1	-1	-1	+1	+1	+1
-1	+1	-1	+1	+1	+1	+1	+1
-1	-1	-1	-1	+1	-1	+1	+1
+1	-1	+1	+1	+1	-1	+1	-1
-1	+1	-1	+1	-1	-1	-1	+1
-1	-1	-1	-1	-1	-1	-1	-1
-2.34	0	0	0	0	0	0	0
2.34	0	0	0	0	0	0	0
0	-2.34	0	0	0	0	0	0
0	2.34	0	0	0	0	0	0
0	0	-2.34	0	0	0	0	0
0	0	2.34	0	0	0	0	0
0	0	0	-2.34	0	0	0	0
0	0	0	2.34	0	0	0	0
0	0	0	0	-2.34	0	0	0
0	0	0	0	2.34	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

12.21. Consider the alloy cracking experiment in Problem 6.15. Suppose that temperature (A) is a noise variable. Find the response model, and the model for the mean response, and the model for the transmitted variability. Can you find settings for the controllable factors that minimize crack length and make the transmitted variability small?

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

Design Expert Output

Response 1 Crack Length
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

Source	**Sum of Squares**	**DF**	**Mean Square**	**F Value**	**Prob > F**	
Model	570.74	8	71.34	1092.46	< 0.0001	significant
A-z1	*72.91*	*1*	*72.91*	*1116.43*	*< 0.0001*	
B-x1	*126.46*	*1*	*126.46*	*1936.46*	*< 0.0001*	
C-x2	*103.46*	*1*	*103.46*	*1584.32*	*< 0.0001*	
D-x3	*30.66*	*1*	*30.66*	*469.52*	*< 0.0001*	
AB	*29.93*	*1*	*29.93*	*458.26*	*< 0.0001*	
AC	*128.50*	*1*	*128.50*	*1967.63*	*< 0.0001*	
BC	*0.074*	*1*	*0.074*	*1.13*	*0.2990*	
ABC	*78.75*	*1*	*78.75*	*1205.90*	*< 0.0001*	
Residual	1.50	23	0.065			
Lack of Fit	*0.20*	*7*	*0.029*	*0.36*	*0.9135*	*not significant*
Pure Error	*1.30*	*16*	*0.081*			
Cor Total	572.25	31				

Design Expert Output

Final Equation in Terms of Coded Factors:

Crack Length=
+11.99
+1.51 *A (z1)
+1.99 *B (x1)
-1.80 *C (x2)
+0.98 *D (x3)

+0.97	*A*B	(z1*x1)
-2.00	*A*C	(z1*x2)
+0.048	*B*C	(x1*x2)
+1.57	*A* B*C	(z1*x1*x2)

From the analysis above, the response model is:

$$\begin{aligned}\hat{y}(\mathbf{x},z_1) &= \hat{\beta}_0+\hat{\beta}_1x_1+\hat{\beta}_2x_2+\hat{\beta}_3x_3+\hat{\beta}_{12}x_1x_2+\hat{\gamma}_1z_1+\hat{\delta}_{11}x_1z_1+\hat{\delta}_{21}x_2z_1+\hat{\delta}_{121}x_1x_2z_1\\ &= 11.99+1.99x_1-1.80x_2+0.98x_3+\\ &\quad +0.048x_1x_2+1.51z_1+0.97x_1z_1-2.00x_2z_1+1.57x_1x_2z_1\end{aligned}$$

The mean model is:

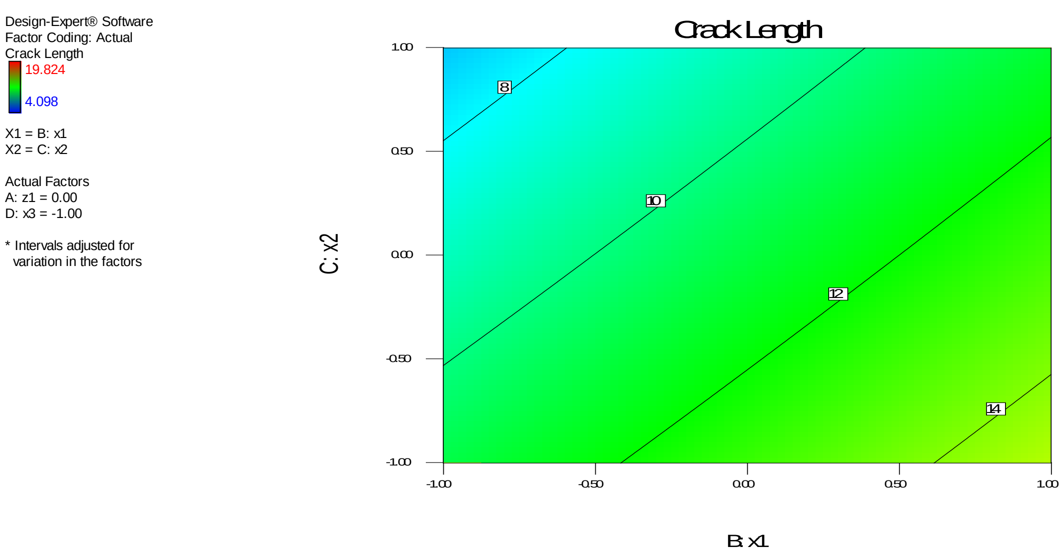
$$\begin{aligned}E_z\left[y(\mathbf{x},z_1)\right] &= \hat{\beta}_0+\hat{\beta}_1x_1+\hat{\beta}_2x_2+\hat{\beta}_3x_3+\hat{\beta}_{12}x_1x_2\\ &= 11.99+1.99x_1-1.80x_2+0.98x_3+0.048x_1x_2\end{aligned}$$

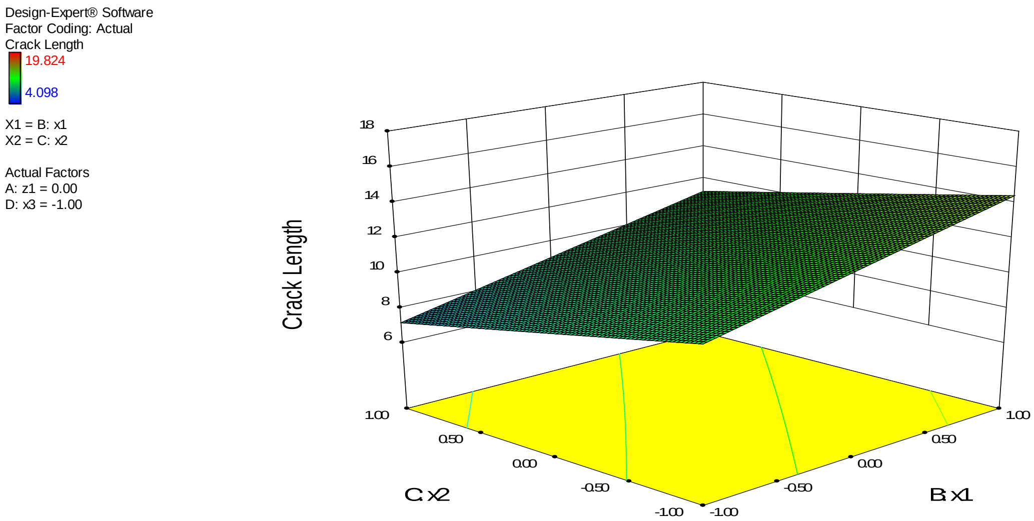
The variance model is shown below. Assuming the low and high levels of the noise variable have been run at one standard deviation on either side of the average value, $\sigma_z^2=1$, and $\sigma^2=0.065$, the residual mean square as calculated above.

$$\begin{aligned}V_z\left[y(\mathbf{x},z_1)\right] &= \sigma_z^2\left(\hat{\gamma}_1+\hat{\delta}_{11}x_1+\hat{\delta}_{21}x_2+\hat{\delta}_{121}x_1x_2\right)^2+\sigma^2\\ &= \hat{\gamma}_1^2+2\hat{\gamma}_1\hat{\delta}_{11}x_1+2\hat{\gamma}_1\hat{\delta}_{21}x_2+2\hat{\gamma}_1\hat{\delta}_{121}x_1x_2+\hat{\delta}_{11}^2x_1^2+\\ &\quad +2\hat{\delta}_{11}\hat{\delta}_{21}x_1x_2+2\hat{\delta}_{11}\hat{\delta}_{121}x_1^2x_2+\hat{\delta}_{21}^2x_2^2+2\hat{\delta}_{21}\hat{\delta}_{121}x_1x_2^2+\hat{\delta}_{121}^2x_1^2x_2^2\end{aligned}$$

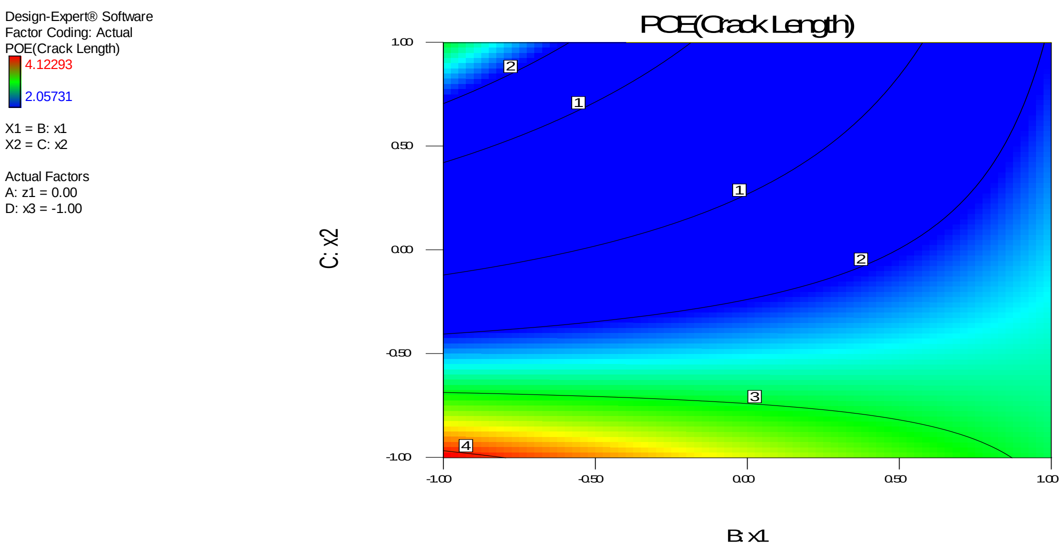
$$\begin{aligned}V_z\left[y(\mathbf{x},z_1)\right] &= (1.51+0.97x_1-2.00x_2+1.57x_1x_2)^2+0.065\\ &= 2.35+2.93x_1-6.04x_2+0.86x_1x_2+0.94x_1^2+\\ &\quad +3.05x_1^2x_2+4.00x_2^2-6.28x_1x_2^2+2.46x_1^2x_2^2\end{aligned}$$

The contour and 3-D plots for the mean cracking model is shown below. Factor x_3 (*D*) is set at the -1 level as this produces the smallest crack length.

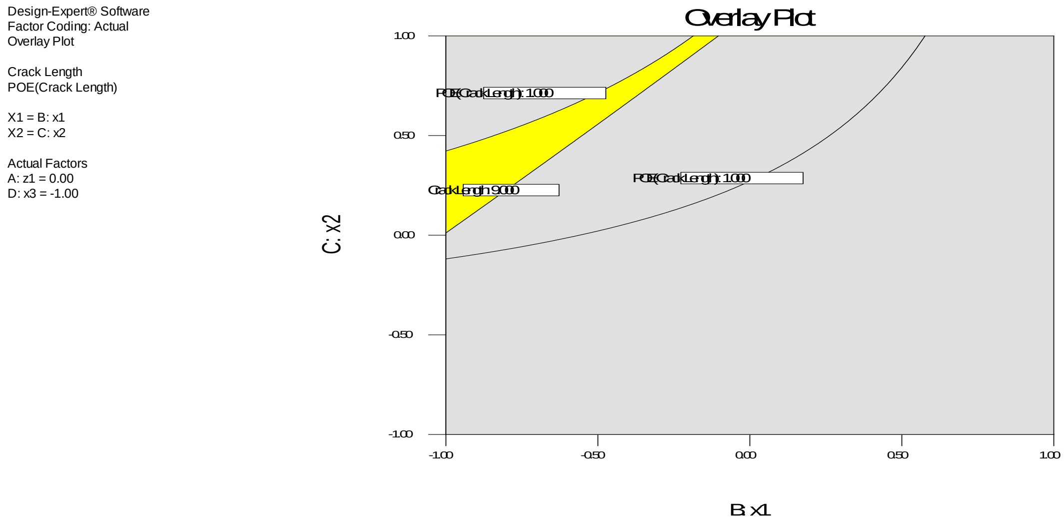




The contour plot of the POE from Design Expert is shown below. Factor x_3 (D) is not included in the variance model and remains at the -1 level for this POE contour plot and the overlay plot.



The overlay plot of the mean and POE contour plots below identify satisfactory operating conditions for the process mean and variance, Crack Length ≤ 9μm, and POE ≤ 1.



Chapter 13
Experiments with Random Factors
Solutions

13.1. An experiment was performed to investigate the capability of a measurement system. Ten parts were randomly selected, and two randomly selected operators measured each part three times. The tests were made in random order, and the data are shown in Table P13.1.

Table P13.1

	Operator 1			Operator 2		
Part	Measurements			Measurements		
Number	1	2	3	1	2	3
1	50	49	50	50	48	51
2	52	52	51	51	51	51
3	53	50	50	54	52	51
4	49	51	50	48	50	51
5	48	49	48	48	49	48
6	52	50	50	52	50	50
7	51	51	51	51	50	50
8	52	50	49	53	48	50
9	50	51	50	51	48	49
10	47	46	49	46	47	48

(a) Analyze the data from this experiment.

Minitab Output

ANOVA: Measurement versus Part, Operator

```
Factor     Type Levels Values
Part     random     10      1      2      3      4      5      6      7
                            8      9     10
Operator random      2      1      2

Analysis of Variance for Measurem

Source            DF         SS         MS       F      P
Part               9     99.017     11.002   18.28  0.000
Operator           1      0.417      0.417    0.69  0.427
Part*Operator      9      5.417      0.602    0.40  0.927
Error             40     60.000      1.500
Total             59    164.850

Source           Variance Error Expected Mean Square for Each Term
                component term (using restricted model)
 1 Part           1.73333   3   (4) + 3(3) + 6(1)
 2 Operator      -0.00617   3   (4) + 3(3) + 30(2)
 3 Part*Operator -0.29938   4   (4) + 3(3)
 4 Error          1.50000       (4)
```

(b) Find point estimates of the variance components using the analysis of variance method.

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 1.5$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}^2_{\tau\beta} = \frac{0.6018519 - 1.5000000}{3} < 0 \text{, assume } \hat{\sigma}^2_{\tau\beta} = 0$$

$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_{AB}}{an} \qquad \hat{\sigma}_\beta^2 = \frac{11.001852 - 0.6018519}{2(3)} = 1.7333$$

$$\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{AB}}{bn} \qquad \hat{\sigma}_\tau^2 = \frac{0.416667 - 0.6018519}{10(3)} < 0 \text{, assume } \hat{\sigma}_\tau^2 = 0$$

All estimates agree with the *Minitab* output.

13.2. An article by Hoof and Berman ("Statistical Analysis of Power Module Thermal Test Equipment Performance", *IEEE Transactions on Components, Hybrids, and Manufacturing Technology* Vol. 11, pp. 516-520, 1988) describes an experiment conducted to investigate the capability of measurements on thermal impedance (C°/W x 100) on a power module for an induction motor starter. There are 10 parts, three operators, and three replicates. The data are shown in Table P13.2.

Table P13.2

Part	Inspector 1			Inspector 2			Inspector 3		
Number	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
1	37	38	37	41	41	40	41	42	41
2	42	41	43	42	42	42	43	42	43
3	30	31	31	31	31	31	29	30	28
4	42	43	42	43	43	43	42	42	42
5	28	30	29	29	30	29	31	29	29
6	42	42	43	45	45	45	44	46	45
7	25	26	27	28	28	30	29	27	27
8	40	40	40	43	42	42	43	43	41
9	25	25	25	27	29	28	26	26	26
10	35	34	34	35	35	34	35	34	35

(a) Analyze the data from this experiment, assuming both parts and operators are random effects.

Minitab Output

ANOVA: Impedance versus Inspector, Part

```
Factor     Type Levels Values
Inspecto random      3     1     2     3
Part     random     10     1     2     3     4     5     6     7
                           8     9    10


Analysis of Variance for Impedanc

Source           DF         SS         MS       F      P
Inspecto          2      39.27      19.63    7.28  0.005
Part              9    3935.96     437.33  162.27  0.000
Inspecto*Part    18      48.51       2.70    5.27  0.000
Error            60      30.67       0.51
Total            89    4054.40


Source           Variance Error Expected Mean Square for Each Term
                component term (using restricted model)
 1 Inspecto        0.5646   3   (4) + 3(3) + 30(1)
 2 Part           48.2926   3   (4) + 3(3) + 9(2)
 3 Inspecto*Part   0.7280   4   (4) + 3(3)
 4 Error           0.5111       (4)
```

(b) Estimate the variance components using the analysis of variance method.

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 0.51$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}^2_{\tau\beta} = \frac{2.70 - 0.51}{3} = 0.73$$

$$\hat{\sigma}^2_{\beta} = \frac{MS_B - MS_{AB}}{an} \qquad \hat{\sigma}^2_{\beta} = \frac{437.33 - 2.70}{3(3)} = 48.29$$

$$\hat{\sigma}^2_{\tau} = \frac{MS_A - MS_{AB}}{bn} \qquad \hat{\sigma}^2_{\tau} = \frac{19.63 - 2.70}{10(3)} = 0.56$$

All estimates agree with the *Minitab* output.

13.3. Reconsider the data in Problem 5.8. Suppose that both factors, machines and operators, are chosen at random.

(a) Analyze the data from this experiment.

	Machine			
Operator	1	2	3	4
1	109	110	108	110
	110	115	109	108
2	110	110	111	114
	112	111	109	112
3	116	112	114	120
	114	115	119	117

The following *Minitab* output contains the analysis of variance and the variance component estimates:

Minitab Output

ANOVA: Strength versus Operator, Machine

```
Factor     Type Levels Values
Operator random      3     1     2     3
Machine  random      4     1     2     3     4

Analysis of Variance for Strength

Source              DF         SS         MS       F      P
Operator             2    160.333     80.167   10.77  0.010
Machine              3     12.458      4.153    0.56  0.662
Operator*Machine     6     44.667      7.444    1.96  0.151
Error               12     45.500      3.792
Total               23    262.958

Source              Variance Error Expected Mean Square for Each Term
                   component term (using restricted model)
 1 Operator           9.0903   3   (4) + 2(3) + 8(1)
 2 Machine           -0.5486   3   (4) + 2(3) + 6(2)
 3 Operator*Machine   1.8264   4   (4) + 2(3)
 4 Error              3.7917       (4)
```

(b) Find point estimates of the variance components using the analysis of variance method.

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 3.79167$$

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}_{\tau\beta}^2 = \frac{7.44444 - 3.79167}{2} = 1.82639$$

$$\hat{\sigma}_{\beta}^2 = \frac{MS_B - MS_{AB}}{an} \qquad \hat{\sigma}_{\beta}^2 = \frac{4.15278 - 7.44444}{3(2)} < 0, \text{ assume } \hat{\sigma}_{\beta}^2 = 0$$

$$\hat{\sigma}_{\tau}^2 = \frac{MS_A - MS_{AB}}{bn} \qquad \hat{\sigma}_{\tau}^2 = \frac{80.16667 - 7.44444}{4(2)} = 9.09028$$

These results agree with the *Minitab* variance component analysis.

13.4. Reconsider the data in Problem 5.15. Suppose that both factors are random.

(a) Analyze the data from this experiment.

	Column	Factor		
Row Factor	1	2	3	4
1	36	39	36	32
2	18	20	22	20
3	30	37	33	34

Minitab Output

General Linear Model: Response versus Row, Column

```
Factor     Type Levels Values
Row      random      3 1 2 3
Column   random      4 1 2 3 4

Analysis of Variance for Response, using Adjusted SS for Tests

Source       DF     Seq SS     Adj SS     Adj MS       F      P
Row           2    580.500    580.500    290.250   60.40     **
Column        3     28.917     28.917      9.639    2.01     **
Row*Column    6     28.833     28.833      4.806      **
Error         0      0.000      0.000      0.000
Total        11    638.250

** Denominator of F-test is zero.

Expected Mean Squares, using Adjusted SS

Source          Expected Mean Square for Each Term
 1 Row          (4) +        (3) +  4.0000(1)
 2 Column       (4) +        (3) +  3.0000(2)
 3 Row*Column   (4) +        (3)
 4 Error        (4)

Error Terms for Tests, using Adjusted SS

Source          Error DF  Error MS  Synthesis of Error MS
 1 Row                 *     4.806  (3)
 2 Column              *     4.806  (3)
 3 Row*Column          *         *  (4)

Variance Components, using Adjusted SS

Source      Estimated Value
Row                 71.3611
Column               1.6111
Row*Column           4.8056
Error                0.0000
```

(b) Estimate the variance components.

Because the experiment is unreplicated and the interaction term was included in the model, there is no estimate of MS_E, and therefore, no estimate of σ^2.

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}_{\tau\beta}^2 = \frac{4.8056 - 0}{1} = 4.8056$$
$$\hat{\sigma}_{\beta}^2 = \frac{MS_B - MS_{AB}}{an} \qquad \hat{\sigma}_{\beta}^2 = \frac{9.6389 - 4.8056}{3(1)} = 1.6111$$
$$\hat{\sigma}_{\tau}^2 = \frac{MS_A - MS_{AB}}{bn} \qquad \hat{\sigma}_{\tau}^2 = \frac{290.2500 - 4.8056}{4(1)} = 71.3611$$

These estimates agree with the *Minitab* output.

13.5. Suppose that in Problem 5.13 the furnace positions were randomly selected, resulting in a mixed model experiment. Reanalyze the data from this experiment under this new assumption. Estimate the appropriate model components.

	Temperature (°C)		
Position	800	825	850
	570	1063	565
1	565	1080	510
	583	1043	590
	528	988	526
2	547	1026	538
	521	1004	532

The following analysis assumes a restricted model:

Minitab Output

ANOVA: Density versus Position, Temperature

```
Factor     Type Levels Values
Position random      2     1     2
Temperat  fixed      3   800   825   850

Analysis of Variance for Density

Source               DF         SS         MS       F      P
Position              1       7160       7160   16.00  0.002
Temperat              2     945342     472671 1155.52  0.001
Position*Temperat     2        818        409    0.91  0.427
Error                12       5371        448
Total                17     958691

Source               Variance Error Expected Mean Square for Each Term
                    component term (using restricted model)
 1 Position            745.83   4   (4) + 9(1)
 2 Temperat                     3   (4) + 3(3) + 6Q[2]
 3 Position*Temperat   -12.83   4   (4) + 3(3)
 4 Error               447.56       (4)
```

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 447.56$$

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}_{\tau\beta}^2 = \frac{409 - 448}{3} < 0 \text{ assume } \hat{\sigma}_{\tau\beta}^2 = 0$$

$$\hat{\sigma}_{\tau}^2 = \frac{MS_A - MS_E}{bn} \qquad \hat{\sigma}_{\tau}^2 = \frac{7160 - 448}{3(3)} = 745.83$$

These results agree with the *Minitab* output.

13.6. Reanalyze the measurement systems experiment in Problem 13.1, assuming that operators are a fixed factor. Estimate the appropriate model components.

Table P13.1

	Operator 1			Operator 2		
	Measurements			Measurements		
Part Number	1	2	3	1	2	3
1	50	49	50	50	48	51
2	52	52	51	51	51	51
3	53	50	50	54	52	51
4	49	51	50	48	50	51
5	48	49	48	48	49	48
6	52	50	50	52	50	50
7	51	51	51	51	50	50
8	52	50	49	53	48	50
9	50	51	50	51	48	49
10	47	46	49	46	47	48

The following analysis assumes a restricted model:

Minitab Output

ANOVA: Measurement versus Part, Operator

```
Factor     Type Levels Values
Part     random     10     1     2     3     4     5     6     7
                           8     9    10
Operator  fixed      2     1     2

Analysis of Variance for Measurem

Source           DF        SS        MS      F      P
Part              9    99.017    11.002   7.33  0.000
Operator          1     0.417     0.417   0.69  0.427
Part*Operator     9     5.417     0.602   0.40  0.927
Error            40    60.000     1.500
Total            59   164.850

Source           Variance Error Expected Mean Square for Each Term
                component term (using restricted model)
 1 Part            1.5836   4   (4) + 6(1)
 2 Operator                 3   (4) + 3(3) + 30Q[2]
 3 Part*Operator  -0.2994   4   (4) + 3(3)
 4 Error           1.5000       (4)
```

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 1.5000$$

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}_{\tau\beta}^2 = \frac{0.60185 - 1.5000}{3} < 0 \text{ assume } \hat{\sigma}_{\tau\beta}^2 = 0$$

$$\hat{\sigma}_{\tau}^{2} = \frac{MS_A - MS_E}{bn} \qquad \hat{\sigma}_{\tau}^{2} = \frac{11.00185 - 1.50000}{2(3)} = 1.58364$$

These results agree with the *Minitab* output.

13.7. Reanalyze the measurement system experiment in Problem 13.2, assuming that operators are a fixed factor. Estimate the appropriate model components.

Table P13.2

Part	Inspector 1			Inspector 2			Inspector 3		
Number	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
1	37	38	37	41	41	40	41	42	41
2	42	41	43	42	42	42	43	42	43
3	30	31	31	31	31	31	29	30	28
4	42	43	42	43	43	43	42	42	42
5	28	30	29	29	30	29	31	29	29
6	42	42	43	45	45	45	44	46	45
7	25	26	27	28	28	30	29	27	27
8	40	40	40	43	42	42	43	43	41
9	25	25	25	27	29	28	26	26	26
10	35	34	34	35	35	34	35	34	35

Minitab Output

ANOVA: Impedance versus Inspector, Part

```
Factor     Type Levels Values
Inspecto  fixed      3     1     2     3
Part     random     10     1     2     3     4     5     6     7
                           8     9    10

Analysis of Variance for Impedanc

Source           DF         SS         MS       F      P
Inspecto          2      39.27      19.63    7.28  0.005
Part              9    3935.96     437.33  855.64  0.000
Inspecto*Part    18      48.51       2.70    5.27  0.000
Error            60      30.67       0.51
Total            89    4054.40

Source           Variance Error Expected Mean Square for Each Term
                component term (using restricted model)
 1 Inspecto                3   (4) + 3(3) + 30Q[1]
 2 Part            48.5353   4   (4) + 9(2)
 3 Inspecto*Part    0.7280   4   (4) + 3(3)
 4 Error            0.5111       (4)
```

$$\hat{\sigma}^{2} = MS_E \qquad \hat{\sigma}^{2} = 0.51$$

$$\hat{\sigma}_{\tau\beta}^{2} = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}_{\tau\beta}^{2} = \frac{2.70 - 0.51}{3} = 0.73$$

$$\hat{\sigma}_{\beta}^{2} = \frac{MS_B - MS_E}{an} \qquad \hat{\sigma}_{\beta}^{2} = \frac{437.33 - 0.51}{3(3)} = 48.54$$

These results agree with the *Minitab* output.

13.8. In problem 5.8, suppose that there are only four machines of interest, but the operators were selected at random.

Operator	Machine			
	1	2	3	4
1	109	110	108	110
	110	115	109	108
2	110	110	111	114
	112	111	109	112
3	116	112	114	120
	114	115	119	117

(a) What type of model is appropriate?

A mixed model is appropriate.

(b) Perform the analysis and estimate the model components.

The following analysis assumes a restricted model:

Minitab Output

```
ANOVA: Strength versus Operator, Machine

Factor     Type Levels Values
Operator random      3     1     2     3
Machine   fixed      4     1     2     3     4

Analysis of Variance for Strength

Source              DF         SS         MS       F      P
Operator             2    160.333     80.167   21.14  0.000
Machine              3     12.458      4.153    0.56  0.662
Operator*Machine     6     44.667      7.444    1.96  0.151
Error               12     45.500      3.792
Total               23    262.958

Source              Variance Error Expected Mean Square for Each Term
                   component term (using restricted model)
 1 Operator            9.547   4   (4) + 8(1)
 2 Machine                     3   (4) + 2(3) + 6Q[2]
 3 Operator*Machine    1.826   4   (4) + 2(3)
 4 Error               3.792       (4)
```

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 3.792$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}^2_{\tau\beta} = \frac{7.444 - 3.792}{2} = 1.826$$

$$\hat{\sigma}^2_{\tau} = \frac{MS_A - MS_E}{bn} \qquad \hat{\sigma}^2_{\tau} = \frac{80.167 - 3.792}{4(2)} = 9.547$$

These results agree with the *Minitab* output.

13.9. Rework Problem 13.5 using the REML method.

	Temperature (°C)		
Position	800	825	850
	570	1063	565
1	565	1080	510
	583	1043	590
	528	988	526
2	547	1026	538
	521	1004	532

The JMP REML output is shown below. The variance components are similar to those calculated in Problem 13.5.

JMP Output

RSquare	0.993347
RSquare Adj	0.99246
Root Mean Square Error	21.15551
Mean of Response	709.9444
Observations (or Sum Wgts)	18

Parameter Estimates

Term	**Estimate**	**Std Error**	**DFDen**	**t Ratio**	**Prob>|t|**
Intercept	709.94444	19.94444	1	35.60	0.0179*
Temperature[800]	-157.6111	6.741707	2	-23.38	0.0018*
Temperature[825]	324.05556	6.741707	2	48.07	0.0004*

REML Variance Component Estimates

Random Effect	**Var Ratio**	**Var Component**	**Std Error**	**95% Lower**	**95% Upper**	**Pct of Total**
Position	1.6760179	750.11111	1126.0118	-1456.832	2957.0538	63.309
Position*Temperature	-0.028674	-12.83333	149.33585	-305.5262	279.85956	-1.083
Residual		447.55556	182.71379	230.13858	1219.556	37.774
Total		1184.8333				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	**Position**	**Position*Temperature**	**Residual**
Position	1267902.7	-6197.276	-1.412e-9
Position*Temperature	-6197.276	22301.197	-11128.11
Residual	-1.412e-9	-11128.11	33384.329

Fixed Effect Tests

Source	**Nparm**	**DF**	**DFDen**	**F Ratio**	**Prob > F**
Temperature	2	2	2	1155.518	0.0009*

13.10. Rework Problem 13.6 using the REML method.

Table P13.1

Part Number	Operator 1 Measurements 1	2	3	Operator 2 Measurements 1	2	3
1	50	49	50	50	48	51
2	52	52	51	51	51	51
3	53	50	50	54	52	51
4	49	51	50	48	50	51
5	48	49	48	48	49	48
6	52	50	50	52	50	50
7	51	51	51	51	50	50
8	52	50	49	53	48	50
9	50	51	50	51	48	49
10	47	46	49	46	47	48

The JMP REML output is shown below. The variance components are similar to those calculated in Problem 13.6.

JMP Output

RSquare	0.420766
RSquare Adj	0.410779
Root Mean Square Error	1.224745
Mean of Response	49.95
Observations (or Sum Wgts)	60

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	49.95	0.42821	9	116.65	<.0001*
Operator[Operator 1]	0.0833333	0.100154	9	0.83	0.4269

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Part Number	1.1555556	1.7333333	0.8656795	0.0366326	3.430034	59.078
Part Number*Operator	-0.199588	-0.299383	0.1464372	-0.586394	-0.012371	-10.204
Residual		1.5	0.3354102	1.0110933	2.4556912	51.126
Total		2.9339506				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	Part Number	Part Number*Operator	Residual
Part Number	0.749401	-0.004472	-1.43e-14
Part Number*Operator	-0.004472	0.0214438	-0.0375
Residual	-1.43e-14	-0.0375	0.1125

Fixed Effect Tests

Source	Nparm	DF	DFDen	F Ratio	Prob > F
Operator	1	1	9	0.6923	0.4269

13.11. Rework Problem 13.7 using the REML method.

Table P13.2

Part Number	Inspector 1 Test 1	Test 2	Test 3	Inspector 2 Test 1	Test 2	Test 3	Inspector 3 Test 1	Test 2	Test 3
1	37	38	37	41	41	40	41	42	41
2	42	41	43	42	42	42	43	42	43
3	30	31	31	31	31	31	29	30	28
4	42	43	42	43	43	43	42	42	42

5	28	30	29	29	30	29	31	29	29
6	42	42	43	45	45	45	44	46	45
7	25	26	27	28	28	30	29	27	27
8	40	40	40	43	42	42	43	43	41
9	25	25	25	27	29	28	26	26	26
10	35	34	34	35	35	34	35	34	35

The JMP REML output is shown below. The variance components are similar to those calculated in Problem 13.7.

JMP Output

RSquare	0.992005
RSquare Adj	0.991821
Root Mean Square Error	0.71492
Mean of Response	35.8
Observations (or Sum Wgts)	90

Parameter Estimates

Term	**Estimate**	**Std Error**	**DFDen**	**t Ratio**	**Prob>|t|**
Intercept	35.8	2.20436	9	16.24	<.0001*
Inspector[Inspector 1]	-0.9	0.244725	18	-3.68	0.0017*
Inspector[Inspector 2]	0.6666667	0.244725	18	2.72	0.0139*

REML Variance Component Estimates

Random Effect	**Var Ratio**	**Var Component**	**Std Error**	**95% Lower**	**95% Upper**	**Pct of Total**
Part Number	94.485507	48.292593	22.906727	3.3962334	93.188952	97.498
Part Number*Inspector	1.4243156	0.7279835	0.3010625	0.1379119	1.3180552	1.470
Residual		0.5111111	0.0933157	0.3681575	0.757543	1.032
Total		49.531687				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	**Part Number**	**Part Number*Inspector**	**Residual**
Part Number	524.71813	-0.02989	-2.76e-13
Part Number*Inspector	-0.02989	0.0906386	-0.002903
Residual	-2.76e-13	-0.002903	0.0087078

Fixed Effect Tests

Source	**Nparm**	**DF**	**DFDen**	**F Ratio**	**Prob > F**
Inspector	2	2	18	7.2849	0.0048*

13.12. Rework Problem 13.8 using the REML method.

Operator	Machine			
	1	2	3	4
1	109	110	108	110
	110	115	109	108
2	110	110	111	114
	112	111	109	112
3	116	112	114	120
	114	115	119	117

The JMP REML output is shown below. The variance components are similar to those calculated in Problem 13.8.

JMP Output

RSquare	0.78154
RSquare Adj	0.748771
Root Mean Square Error	1.94722
Mean of Response	112.2917
Observations (or Sum Wgts)	24

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	112.29167	1.827643	2	61.44	0.0003*
Machine[1]	-0.458333	0.964653	6	-0.48	0.6515
Machine[2]	-0.125	0.964653	6	-0.13	0.9011
Machine[3]	-0.625	0.964653	6	-0.65	0.5410

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Operator	2.3974359	9.0902778	10.035225	-10.5784	28.758958	61.804
Operator*Machine	0.481685	1.8263889	2.2841505	-2.650464	6.3032416	12.417
Residual		3.7916667	1.5479414	1.9497217	10.332013	25.779
Total		14.708333				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	Operator	Operator*Machine	Residual
Operator	100.70575	-1.154578	1.686e-12
Operator*Machine	-1.154578	5.2173434	-1.198061
Residual	1.686e-12	-1.198061	2.3961227

Fixed Effect Tests

Source	Nparm	DF	DFDen	F Ratio	Prob > F
Machine	3	3	6	0.5578	0.6619

13.13. By application of the expectation operator, develop the expected mean squares for the two-factor factorial, mixed model. Use the restricted model assumptions. Check your results with the expected mean squares given in Equation 13.9 to see that they agree.

The sums of squares may be written as

$$SS_A = bn\sum_{i=1}^{a}(\bar{y}_{i..} - \bar{y}_{...})^2 , \quad SS_B = an\sum_{j=1}^{b}(\bar{y}_{.j.} - \bar{y}_{...})^2$$

$$SS_{AB} = n\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 , \quad SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{...})^2$$

Using the model $y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk}$, we may find that

$$\bar{y}_{i..} = \mu + \tau_i + (\tau\beta)_{i.} + \bar{\varepsilon}_{i..}$$
$$\bar{y}_{.j.} = \mu + \beta_j + \bar{\varepsilon}_{.j.}$$
$$\bar{y}_{ij.} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \bar{\varepsilon}_{ij.}$$
$$\bar{y}_{...} = \mu + \beta_. + \bar{\varepsilon}_{...}$$

Using the assumptions for the restricted form of the mixed model, $\tau_. = 0$, $(\tau\beta)_{.j} = 0$, which imply that $(\tau\beta)_{..} = 0$. Substituting these expressions into the sums of squares yields

$$SS_A = bn\sum_{i=1}^{a}\left(\tau_i + (\tau\beta)_{i.} + \bar{\varepsilon}_{i..} - \bar{\varepsilon}_{...}\right)^2$$
$$SS_B = an\sum_{j=1}^{b}\left(\beta_j + \bar{\varepsilon}_{.j.} - \bar{\varepsilon}_{...}\right)^2$$
$$SS_{AB} = n\sum_{i=1}^{a}\sum_{j=1}^{b}\left((\tau\beta)_{ij} - (\tau\beta)_{i.} + \bar{\varepsilon}_{ij.} - \bar{\varepsilon}_{i..} - \bar{\varepsilon}_{.j.} + \bar{\varepsilon}_{...}\right)^2$$
$$SS_E = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\varepsilon_{ijk} - \bar{\varepsilon}_{ij.}\right)^2$$

Using the assumption that $E(\varepsilon_{ijk}) = 0$, $V(\varepsilon_{ijk}) = 0$, and $E(\varepsilon_{ijk} \cdot \varepsilon_{i'j'k'}) = 0$, we may divide each sum of squares by its degrees of freedom and take the expectation to produce

$$E(MS_A) = \sigma^2 + \left[\frac{bn}{(a-1)}\right]E\sum_{i=1}^{a}\left(\tau_i + (\tau\beta)_{i.}\right)^2$$
$$E(MS_B) = \sigma^2 + \left[\frac{an}{(b-1)}\right]\sum_{j=1}^{b}\beta_j^2$$
$$E(MS_{AB}) = \sigma^2 + \left[\frac{n}{(a-1)(b-1)}\right]E\sum_{i=1}^{a}\sum_{j=1}^{b}\left((\tau\beta)_{ij} - (\tau\beta)_{i.}\right)^2$$
$$E(MS_E) = \sigma^2$$

Note that $E(MS_B)$ and $E(MS_E)$ are the results given in Table 8-3. We need to simplify $E(MS_A)$ and $E(MS_{AB})$. Consider $E(MS_A)$

$$E(MS_A) = \sigma^2 + \frac{bn}{a-1}\left[\sum_{i=1}^{a}E(\tau_i)^2 + \sum_{i=1}^{a}E(\tau\beta)_{i.}^2 + (crossproducts = 0)\right]$$
$$E(MS_A) = \sigma^2 + \frac{bn}{a-1}\left[\sum_{i=1}^{a}\tau_i^2 + a\frac{\left[\frac{(a-1)}{a}\right]}{b}\sigma_{\tau\beta}^2\right]$$
$$E(MS_A) = \sigma^2 + n\sigma_{\tau\beta}^2 + \frac{bn}{a-1}\sum_{i=1}^{a}\tau_i^2$$

since $(\tau\beta)_{ij}$ is $NID\left(0, \frac{a-1}{a}\sigma_{\tau\beta}^2\right)$. Consider $E(MS_{AB})$

$$E(MS_{AB}) = \sigma^2 + \frac{n}{(a-1)(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}E\left((\tau\beta)_{ij} - (\tau\beta)_{i.}\right)^2$$
$$E(MS_{AB}) = \sigma^2 + \frac{n}{(a-1)(b-1)}\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\frac{b-1}{b}\right)\left(\frac{a-1}{a}\right)\sigma_{\tau\beta}^2$$
$$E(MS_{AB}) = \sigma^2 + n\sigma_{\tau\beta}^2$$

Thus $E(MS_A)$and $E(MS_{AB})$agree with Equation 13.9.

13.14. Consider the three-factor factorial design in Example 13.5. Propose appropriate test statistics for all main effects and interactions. Repeat for the case where *A* and *B* are fixed and *C* is random.

If all three factors are random there are no exact tests on main effects. We could use the following:

$$A : F = \frac{MS_A + MS_{ABC}}{MS_{AB} + MS_{AC}}$$
$$B : F = \frac{MS_B + MS_{ABC}}{MS_{AB} + MS_{BC}}$$
$$C : F = \frac{MS_C + MS_{ABC}}{MS_{AC} + MS_{BC}}$$

If *A* and *B* are fixed and *C* is random, the expected mean squares are (assuming the restricted form of the model):

Factor	F *a* *i*	F *b* *j*	R *c* *k*	R *n* *l*	E(*MS*)
τ_i	0	*b*	*c*	*n*	$\sigma^2 + bn\sigma^2_{\tau\gamma} + bcn\sum\frac{\tau_i^2}{(a-1)}$
β_j	*a*	0	*c*	*n*	$\sigma^2 + an\sigma^2_{\beta\gamma} + acn\sum\frac{\beta_j^2}{(b-1)}$
γ_k	*a*	*b*	1	*n*	$\sigma^2 + abn\sigma^2_{\gamma}$
$(\tau\beta)_{ij}$	0	0	*c*	*n*	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sum\sum\frac{(\tau\beta)_{ij}^2}{(a-1)(b-1)}$
$(\tau\gamma)_{ik}$	0	*b*	1	*n*	$\sigma^2 + bn\sigma^2_{\tau\gamma}$
$(\beta\gamma)_{jk}$	*a*	0	1	*n*	$\sigma^2 + an\sigma^2_{\beta\gamma}$
$(\tau\beta\gamma)_{ijk}$	0	0	1	*n*	$\sigma^2 + n\sigma^2_{\tau\beta\gamma}$
$\varepsilon_{(ijk)l}$	1	1	1	1	σ^2

These are exact tests for all effects.

13.15. Consider the experiment in Example 13.6. Analyze the data for the case where *A*, *B*, and *C* are random.

Minitab Output

```
ANOVA: Drop versus Temp, Operator, Gauge

Factor      Type Levels Values
Temp      random      3     60     75     90
Operator random      4      1      2      3      4
Gauge     random      3      1      2      3

Analysis of Variance for Drop

Source                  DF         SS          MS        F       P
Temp                     2    1023.36      511.68     2.30   0.171 x
```

```
Operator                3     423.82     141.27    0.63  0.616 x
Gauge                   2       7.19       3.60    0.06  0.938 x
Temp*Operator           6    1211.97     202.00   14.59  0.000
Temp*Gauge              4     137.89      34.47    2.49  0.099
Operator*Gauge          6     209.47      34.91    2.52  0.081
Temp*Operator*Gauge    12     166.11      13.84    0.65  0.788
Error                  36     770.50      21.40
Total                  71    3950.32

x Not an exact F-test.

Source                  Variance Error Expected Mean Square for Each Term
                       component term (using restricted model)
 1 Temp                   12.044   *   (8) + 2(7) + 8(5) + 6(4) + 24(1)
 2 Operator               -4.544   *   (8) + 2(7) + 6(6) + 6(4) + 18(2)
 3 Gauge                  -2.164   *   (8) + 2(7) + 6(6) + 8(5) + 24(3)
 4 Temp*Operator          31.359   7   (8) + 2(7) + 6(4)
 5 Temp*Gauge              2.579   7   (8) + 2(7) + 8(5)
 6 Operator*Gauge          3.512   7   (8) + 2(7) + 6(6)
 7 Temp*Operator*Gauge    -3.780   8   (8) + 2(7)
 8 Error                  21.403       (8)

* Synthesized Test.

Error Terms for Synthesized Tests

Source                   Error DF  Error MS  Synthesis of Error MS
 1 Temp                      6.97    222.63   (4) + (5) - (7)
 2 Operator                  7.09    223.06   (4) + (6) - (7)
 3 Gauge                     5.98     55.54   (5) + (6) - (7)
```

Since all three factors are random there are no exact tests on main effects. *Minitab* uses an approximate *F* test for the these factors.

13.16. Derive the expected mean squares shown in Table 13.11.

Factor	F *a* *i*	R *b* *j*	R *c* *k*	R *n* *l*	E(*MS*)
τ_i	0	b	c	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + bn\sigma^2_{\tau\gamma} + cn\sigma^2_{\tau\beta} + bcn\sum\frac{\tau_i^2}{(a-1)}$
β_j	a	1	c	n	$\sigma^2 + an\sigma^2_{\beta\gamma} + acn\sigma^2_{\beta}$
γ_k	a	b	1	n	$\sigma^2 + an\sigma^2_{\beta\gamma} + abn\sigma^2_{\gamma}$
$(\tau\beta)_{ij}$	0	1	c	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + cn\sigma^2_{\tau\beta}$
$(\tau\gamma)_{ik}$	0	b	1	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma} + bn\sigma^2_{\tau\gamma}$
$(\beta\gamma)_{jk}$	a	1	1	n	$\sigma^2 + an\sigma^2_{\beta\gamma}$
$(\tau\beta\gamma)_{ijk}$	0	1	1	n	$\sigma^2 + n\sigma^2_{\tau\beta\gamma}$
ε_{ijkl}	1	1	1	1	σ^2

13.17. Consider a four-factor factorial experiment where factor *A* is at *a* levels, factor *B* is at *b* levels, factor *C* is at *c* levels, factor *D* is at *d* levels, and there are *n* replicates. Write down the sums of squares, the degrees of freedom, and the expected mean squares for the following cases. Assume the restricted model for all mixed models. You may use a computer package such as *Minitab*. Do exact tests exist for all effects? If not, propose test statistics for those effects that cannot be directly tested.

The four factor model is:

$$y_{ijklh}=\mu+\tau_i+\beta_j+\gamma_k+\delta_l+(\tau\beta)_{ij}+(\tau\gamma)_{ik}+(\tau\delta)_{il}+(\beta\gamma)_{jk}+(\beta\delta)_{jl}+(\gamma\delta)_{kl}+$$
$$(\tau\beta\gamma)_{ijk}+(\tau\beta\delta)_{ijl}+(\beta\gamma\delta)_{jkl}+(\tau\gamma\delta)_{ikl}+(\tau\beta\gamma\delta)_{ijkl}+\varepsilon_{ijklh}$$

To simplify the expected mean square derivations, let capital Latin letters represent the factor effects or variance components. For example, $A=\dfrac{bcdn\sum\tau_i^2}{a-1}$, or $B=acdn\sigma_\beta^2$.

(a) *A*, *B*, *C*, and *D* are fixed factors.

Factor	F *a* *i*	F *b* *j*	F *c* *k*	F *d* *l*	R *n* *h*	E(*MS*)
τ_i	0	b	c	d	n	σ^2+A
β_j	a	0	c	d	n	σ^2+B
γ_k	a	b	0	d	n	σ^2+C
δ_l	a	b	c	0	n	σ^2+D
$(\tau\beta)_{ij}$	0	0	c	d	n	σ^2+AB
$(\tau\gamma)_{ik}$	0	b	0	d	n	σ^2+AC
$(\tau\delta)_{il}$	0	b	c	0	n	σ^2+AD
$(\beta\gamma)_{jk}$	a	0	0	d	n	σ^2+BC
$(\beta\delta)_{jl}$	a	0	c	0	n	σ^2+BD
$(\gamma\delta)_{kl}$	a	b	0	0	n	σ^2+CD
$(\tau\beta\gamma)_{ijk}$	0	0	0	d	n	σ^2+ABC
$(\tau\beta\delta)_{ijl}$	0	0	c	0	n	σ^2+ABD
$(\beta\gamma\delta)_{jkl}$	a	0	0	0	n	σ^2+BCD
$(\tau\gamma\delta)_{ikl}$	0	b	0	0	n	σ^2+ACD
$(\tau\beta\gamma\delta)_{ijkl}$	0	0	0	0	n	σ^2+ABCD
$\varepsilon_{(ijkl)h}$	1	1	1	1	1	σ^2

There are exact tests for all effects. The results can also be generated in *Minitab* as follows:

Minitab Output

```
ANOVA: y versus A, B, C, D
Factor     Type Levels Values
A         fixed      2     H     L
B         fixed      2     H     L
C         fixed      2     H     L
D         fixed      2     H     L

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13    0.49  0.492
B            1       0.13       0.13    0.01  0.921
C            1       1.13       1.13    0.09  0.767
D            1       0.13       0.13    0.01  0.921
A*B          1       3.13       3.13    0.25  0.622
A*C          1       3.13       3.13    0.25  0.622
A*D          1       3.13       3.13    0.25  0.622
B*C          1       3.13       3.13    0.25  0.622
B*D          1       3.13       3.13    0.25  0.622
C*D          1       3.13       3.13    0.25  0.622
A*B*C        1       3.13       3.13    0.25  0.622
A*B*D        1      28.13      28.13    2.27  0.151
```

```
A*C*D        1       3.13       3.13    0.25  0.622
B*C*D        1       3.13       3.13    0.25  0.622
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

Source       Variance Error Expected Mean Square for Each Term
            component term (using restricted model)
 1 A                   16   (16) + 16Q[1]
 2 B                   16   (16) + 16Q[2]
 3 C                   16   (16) + 16Q[3]
 4 D                   16   (16) + 16Q[4]
 5 A*B                 16   (16) + 8Q[5]
 6 A*C                 16   (16) + 8Q[6]
 7 A*D                 16   (16) + 8Q[7]
 8 B*C                 16   (16) + 8Q[8]
 9 B*D                 16   (16) + 8Q[9]
10 C*D                 16   (16) + 8Q[10]
11 A*B*C               16   (16) + 4Q[11]
12 A*B*D               16   (16) + 4Q[12]
13 A*C*D               16   (16) + 4Q[13]
14 B*C*D               16   (16) + 4Q[14]
15 A*B*C*D             16   (16) + 2Q[15]
16 Error      12.38         (16)
```

(b) *A*, *B*, *C*, and *D* are random factors.

	R	R	R	R	R	
	a	*b*	*c*	*d*	*n*	
Factor	*i*	*j*	*k*	*l*	*h*	E(*MS*)
τ_i	1	*b*	*c*	*d*	*n*	$\sigma^2 + ABCD + ACD + ABD + ABC + AD + AC + AB + A$
β_j	*a*	1	*c*	*d*	*n*	$\sigma^2 + ABCD + BCD + ABD + ABC + BD + BC + AB + B$
γ_k	*a*	*b*	1	*d*	*n*	$\sigma^2 + ABCD + ACD + BCD + ABC + AC + BC + CD + C$
δ_l	*a*	*b*	*c*	1	*n*	$\sigma^2 + ABCD + ACD + BCD + ABD + BD + AD + CD + D$
$(\tau\beta)_{ij}$	1	1	*c*	*d*	*n*	$\sigma^2 + ABCD + ABC + ABD + AB$
$(\tau\gamma)_{ik}$	1	*b*	1	*d*	*n*	$\sigma^2 + ABCD + ABC + ACD + AC$
$(\tau\delta)_{il}$	1	*b*	*c*	1	*n*	$\sigma^2 + ABCD + ABD + ACD + AD$
$(\beta\gamma)_{jk}$	*a*	1	1	*d*	*n*	$\sigma^2 + ABCD + ABC + BCD + BC$
$(\beta\delta)_{jl}$	*a*	1	*c*	1	*n*	$\sigma^2 + ABCD + ABD + BCD + BD$
$(\gamma\delta)_{kl}$	*a*	*b*	1	1	*n*	$\sigma^2 + ABCD + ACD + BCD + CD$
$(\tau\beta\gamma)_{ijk}$	1	1	1	*d*	*n*	$\sigma^2 + ABCD + ABC$
$(\tau\beta\delta)_{ijl}$	1	1	*c*	1	*n*	$\sigma^2 + ABCD + ABD$
$(\beta\gamma\delta)_{jkl}$	*a*	1	1	1	*n*	$\sigma^2 + ABCD + BCD$
$(\tau\gamma\delta)_{ikl}$	1	*b*	1	1	*n*	$\sigma^2 + ABCD + ACD$
$(\tau\beta\gamma\delta)_{ijkl}$	1	1	1	1	*n*	$\sigma^2 + ABCD$
$\varepsilon_{(ijkl)h}$	1	1	1	1	1	σ^2

No exact tests exist on main effects or two-factor interactions. For main effects use statistics such as:

$$A{:}F = \frac{MS_A + MS_{ABC} + MS_{ABD} + MS_{ACD}}{MS_{AB} + MS_{AC} + MS_{AD} + MS_{ABCD}}$$

For testing two-factor interactions use statistics such as: $AB{:}F = \dfrac{MS_{AB} + MS_{ABCD}}{MS_{ABC} + MS_{ABD}}$

The results can also be generated in *Minitab* as follows:

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A        random      2     H     L
B        random      2     H     L
C        random      2     H     L
D        random      2     H     L

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13      **
B            1       0.13       0.13      **
C            1       1.13       1.13    0.36  0.843 x
D            1       0.13       0.13      **
A*B          1       3.13       3.13    0.11  0.796 x
A*C          1       3.13       3.13    1.00  0.667 x
A*D          1       3.13       3.13    0.11  0.796 x
B*C          1       3.13       3.13    1.00  0.667 x
B*D          1       3.13       3.13    0.11  0.796 x
C*D          1       3.13       3.13    1.00  0.667 x
A*B*C        1       3.13       3.13    1.00  0.500
A*B*D        1      28.13      28.13    9.00  0.205
A*C*D        1       3.13       3.13    1.00  0.500
B*C*D        1       3.13       3.13    1.00  0.500
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

x Not an exact F-test.

** Denominator of F-test is zero.

Source      Variance Error Expected Mean Square for Each Term
           component term (using restricted model)
 1 A         1.7500   *   (16) + 2(15) + 4(13) + 4(12) + 4(11) + 8(7) + 8(6)
                           + 8(5) + 16(1)
 2 B         1.3750   *   (16) + 2(15) + 4(14) + 4(12) + 4(11) + 8(9) + 8(8)
                           + 8(5) + 16(2)
 3 C        -0.1250   *   (16) + 2(15) + 4(14) + 4(13) + 4(11) + 8(10) + 8(8)
                           + 8(6) + 16(3)
 4 D         1.3750   *   (16) + 2(15) + 4(14) + 4(13) + 4(12) + 8(10) + 8(9)
                           + 8(7) + 16(4)
 5 A*B      -3.1250   *   (16) + 2(15) + 4(12) + 4(11) + 8(5)
 6 A*C       0.0000   *   (16) + 2(15) + 4(13) + 4(11) + 8(6)
 7 A*D      -3.1250   *   (16) + 2(15) + 4(13) + 4(12) + 8(7)
 8 B*C       0.0000   *   (16) + 2(15) + 4(14) + 4(11) + 8(8)
 9 B*D      -3.1250   *   (16) + 2(15) + 4(14) + 4(12) + 8(9)
10 C*D       0.0000   *   (16) + 2(15) + 4(14) + 4(13) + 8(10)
11 A*B*C     0.0000  15   (16) + 2(15) + 4(11)
12 A*B*D     6.2500  15   (16) + 2(15) + 4(12)
13 A*C*D     0.0000  15   (16) + 2(15) + 4(13)
14 B*C*D     0.0000  15   (16) + 2(15) + 4(14)
15 A*B*C*D  -4.6250  16   (16) + 2(15)
16 Error    12.3750       (16)

* Synthesized Test.

Error Terms for Synthesized Tests

Source       Error DF  Error MS  Synthesis of Error MS
 1 A             0.56         *  (5) + (6) + (7) - (11) - (12) - (13) + (15)
 2 B             0.56         *  (5) + (8) + (9) - (11) - (12) - (14) + (15)
 3 C             0.14      3.13  (6) + (8) + (10) - (11) - (13) - (14) + (15)
 4 D             0.56         *  (7) + (9) + (10) - (12) - (13) - (14) + (15)
 5 A*B           0.98     28.13  (11) + (12) - (15)
 6 A*C           0.33      3.13  (11) + (13) - (15)
 7 A*D           0.98     28.13  (12) + (13) - (15)
 8 B*C           0.33      3.13  (11) + (14) - (15)
```

```
 9 B*D          0.98     28.13   (12) + (14) - (15)
10 C*D          0.33      3.13   (13) + (14) - (15)
```

(c) *A* is fixed and *B*, *C*, and *D* are random.

	F	R	R	R	R	
	a	b	c	d	n	
Factor	i	j	k	l	h	E(*MS*)
τ_i	0	b	c	d	n	$\sigma^2 + ABCD + ACD + ABD + ABC + AD + AC + AB + A$
β_j	a	1	c	d	n	$\sigma^2 + BCD + BD + BC + B$
γ_k	a	b	1	d	n	$\sigma^2 + BCD + BC + CD + C$
δ_l	a	b	c	1	n	$\sigma^2 + BCD + BD + CD + D$
$(\tau\beta)_{ij}$	0	1	c	d	n	$\sigma^2 + ABCD + ABC + ABD + AB$
$(\tau\gamma)_{ik}$	0	b	1	d	n	$\sigma^2 + ABCD + ABC + ACD + AC$
$(\tau\delta)_{il}$	0	b	c	1	n	$\sigma^2 + ABCD + ABD + ACD + AD$
$(\beta\gamma)_{jk}$	a	1	1	d	n	$\sigma^2 + BCD + BC$
$(\beta\delta)_{jl}$	a	1	c	1	n	$\sigma^2 + BCD + BD$
$(\gamma\delta)_{kl}$	a	b	1	1	n	$\sigma^2 + BCD + CD$
$(\tau\beta\gamma)_{ijk}$	0	1	1	d	n	$\sigma^2 + ABCD + ABC$
$(\tau\beta\delta)_{ijl}$	0	1	c	1	n	$\sigma^2 + ABCD + ABD$
$(\beta\gamma\delta)_{jkl}$	a	1	1	1	n	$\sigma^2 + BCD$
$(\tau\gamma\delta)_{ikl}$	0	b	1	1	n	$\sigma^2 + ABCD + ACD$
$(\tau\beta\gamma\delta)_{ijkl}$	0	1	1	1	n	$\sigma^2 + ABCD$
$\varepsilon_{(ijkl)h}$	1	1	1	1	1	σ^2

No exact tests exist on main effects or two-factor interactions involving the fixed factor *A*. To test the fixed factor *A* use

$$A\!:F = \frac{MS_A + MS_{ABC} + MS_{ABD} + MS_{ACD}}{MS_{AB} + MS_{AC} + MS_{AD} + MS_{ABCD}}$$

Random main effects could be tested by, for example: $D : F = \dfrac{MS_D + MS_{BCD}}{MS_{BD} + MS_{CD}}$

For testing two-factor interactions involving *A* use: $AB\!:F = \dfrac{MS_{AB} + MS_{ABCD}}{MS_{ABC} + MS_{ABD}}$

The results can also be generated in *Minitab* as follows:

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A         fixed      2     H     L
B        random      2     H     L
C        random      2     H     L
D        random      2     H     L

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13      **
B            1       0.13       0.13    0.04  0.907 x
C            1       1.13       1.13    0.36  0.761 x
D            1       0.13       0.13    0.04  0.907 x
A*B          1       3.13       3.13    0.11  0.796 x
A*C          1       3.13       3.13    1.00  0.667 x
A*D          1       3.13       3.13    0.11  0.796 x
B*C          1       3.13       3.13    1.00  0.500
B*D          1       3.13       3.13    1.00  0.500
C*D          1       3.13       3.13    1.00  0.500
A*B*C        1       3.13       3.13    1.00  0.500
A*B*D        1      28.13      28.13    9.00  0.205
A*C*D        1       3.13       3.13    1.00  0.500
B*C*D        1       3.13       3.13    0.25  0.622
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

x Not an exact F-test.

** Denominator of F-test is zero.

Source      Variance Error Expected Mean Square for Each Term
           component term (using restricted model)
 1 A                  *    (16) + 2(15) + 4(13) + 4(12) + 4(11) + 8(7) + 8(6)
                            + 8(5) + 16Q[1]
 2 B         -0.1875   *   (16) + 4(14) + 8(9) + 8(8) + 16(2)
 3 C         -0.1250   *   (16) + 4(14) + 8(10) + 8(8) + 16(3)
 4 D         -0.1875   *   (16) + 4(14) + 8(10) + 8(9) + 16(4)
 5 A*B       -3.1250   *   (16) + 2(15) + 4(12) + 4(11) + 8(5)
 6 A*C        0.0000   *   (16) + 2(15) + 4(13) + 4(11) + 8(6)
 7 A*D       -3.1250   *   (16) + 2(15) + 4(13) + 4(12) + 8(7)
 8 B*C        0.0000  14   (16) + 4(14) + 8(8)
 9 B*D        0.0000  14   (16) + 4(14) + 8(9)
10 C*D        0.0000  14   (16) + 4(14) + 8(10)
11 A*B*C      0.0000  15   (16) + 2(15) + 4(11)
12 A*B*D      6.2500  15   (16) + 2(15) + 4(12)
13 A*C*D      0.0000  15   (16) + 2(15) + 4(13)
14 B*C*D     -2.3125  16   (16) + 4(14)
15 A*B*C*D   -4.6250  16   (16) + 2(15)
16 Error     12.3750       (16)

* Synthesized Test.

Error Terms for Synthesized Tests

Source       Error DF  Error MS  Synthesis of Error MS
 1 A             0.56         *   (5) + (6) + (7) - (11) - (12) - (13) + (15)
 2 B             0.33      3.13   (8) + (9) - (14)
 3 C             0.33      3.13   (8) + (10) - (14)
 4 D             0.33      3.13   (9) + (10) - (14)
 5 A*B           0.98     28.13   (11) + (12) - (15)
 6 A*C           0.33      3.13   (11) + (13) - (15)
 7 A*D           0.98     28.13   (12) + (13) - (15)
```

(d) *A* and *B* are fixed and *C* and *D* are random.

	F	F	R	R	R	
	a	b	c	d	n	
Factor	i	j	k	l	h	E(*MS*)
τ_i	0	b	c	d	n	$\sigma^2 + ACD + AD + AC + A$
β_j	a	0	c	d	n	$\sigma^2 + BCD + BC + BD + B$
γ_k	a	b	1	d	n	$\sigma^2 + CD + C$
δ_l	a	b	c	1	n	$\sigma^2 + CD + D$
$(\tau\beta)_{ij}$	0	0	c	d	n	$\sigma^2 + ABCD + ABC + ABD + AB$
$(\tau\gamma)_{ik}$	0	b	1	d	n	$\sigma^2 + ACD + AC$
$(\tau\delta)_{il}$	0	b	c	1	n	$\sigma^2 + ACD + AD$
$(\beta\gamma)_{jk}$	a	0	1	d	n	$\sigma^2 + BCD + BC$
$(\beta\delta)_{jl}$	a	0	c	1	n	$\sigma^2 + BCD + BD$
$(\gamma\delta)_{kl}$	a	b	1	1	n	$\sigma^2 + CD$
$(\tau\beta\gamma)_{ijk}$	0	0	1	d	n	$\sigma^2 + ABCD + ABC$
$(\tau\beta\delta)_{ijl}$	0	0	c	1	n	$\sigma^2 + ABCD + ABD$
$(\beta\gamma\delta)_{jkl}$	a	0	1	1	n	$\sigma^2 + BCD$
$(\tau\gamma\delta)_{ikl}$	0	b	1	1	n	$\sigma^2 + ACD$
$(\tau\beta\gamma\delta)_{ijkl}$	0	0	1	1	n	$\sigma^2 + ABCD$
$\varepsilon_{(ijkl)h}$	1	1	1	1	1	σ^2

There are no exact tests on the fixed factors *A* and *B*, or their two-factor interaction *AB*. The appropriate test statistics are:

$$A\colon F = \frac{MS_A + MS_{ACD}}{MS_{AC} + MS_{AD}}$$
$$B\colon F = \frac{MS_B + MS_{BCD}}{MS_{BC} + MS_{BD}}$$
$$AB\colon F = \frac{MS_{AB} + MS_{ABCD}}{MS_{ABC} + MS_{ABD}}$$

The results can also be generated in *Minitab* as follows:

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A         fixed      2     H     L
B         fixed      2     H     L
C        random      2     H     L
D        random      2     H     L

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13    1.96  0.604 x
B            1       0.13       0.13    0.04  0.907 x
C            1       1.13       1.13    0.36  0.656
D            1       0.13       0.13    0.04  0.874
A*B          1       3.13       3.13    0.11  0.796 x
A*C          1       3.13       3.13    1.00  0.500
A*D          1       3.13       3.13    1.00  0.500
B*C          1       3.13       3.13    1.00  0.500
B*D          1       3.13       3.13    1.00  0.500
```

```
C*D          1       3.13       3.13    0.25  0.622
A*B*C        1       3.13       3.13    1.00  0.500
A*B*D        1      28.13      28.13    9.00  0.205
A*C*D        1       3.13       3.13    0.25  0.622
B*C*D        1       3.13       3.13    0.25  0.622
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

x Not an exact F-test.

Source      Variance Error Expected Mean Square for Each Term
           component term (using restricted model)
 1 A                  *   (16) + 4(13) + 8(7) + 8(6) + 16Q[1]
 2 B                  *   (16) + 4(14) + 8(9) + 8(8) + 16Q[2]
 3 C         -0.1250  10   (16) + 8(10) + 16(3)
 4 D         -0.1875  10   (16) + 8(10) + 16(4)
 5 A*B                *   (16) + 2(15) + 4(12) + 4(11) + 8Q[5]
 6 A*C        0.0000  13   (16) + 4(13) + 8(6)
 7 A*D        0.0000  13   (16) + 4(13) + 8(7)
 8 B*C        0.0000  14   (16) + 4(14) + 8(8)
 9 B*D        0.0000  14   (16) + 4(14) + 8(9)
10 C*D       -1.1563  16   (16) + 8(10)
11 A*B*C      0.0000  15   (16) + 2(15) + 4(11)
12 A*B*D      6.2500  15   (16) + 2(15) + 4(12)
13 A*C*D     -2.3125  16   (16) + 4(13)
14 B*C*D     -2.3125  16   (16) + 4(14)
15 A*B*C*D   -4.6250  16   (16) + 2(15)
16 Error     12.3750       (16)

* Synthesized Test.

Error Terms for Synthesized Tests

Source       Error DF  Error MS  Synthesis of Error MS
 1 A             0.33      3.13   (6) + (7) - (13)
 2 B             0.33      3.13   (8) + (9) - (14)
 5 A*B           0.98     28.13   (11) + (12) - (15)
```

(e) *A*, *B* and *C* are fixed and *D* is random.

	F	F	F	R	R	
	a	*b*	*c*	*d*	*n*	
Factor	*i*	*j*	*k*	*l*	*h*	E(*MS*)
τ_i	0	*b*	*c*	*d*	*n*	$\sigma^2 + AD + A$
β_j	*a*	0	*c*	*d*	*n*	$\sigma^2 + BD + B$
γ_k	*a*	*b*	0	*d*	*n*	$\sigma^2 + CD + C$
δ_l	*a*	*b*	*c*	1	*n*	$\sigma^2 + D$
$(\tau\beta)_{ij}$	0	0	*c*	*d*	*n*	$\sigma^2 + ABD + AB$
$(\tau\gamma)_{ik}$	0	*b*	0	*d*	*n*	$\sigma^2 + ACD + AC$
$(\tau\delta)_{il}$	0	*b*	*c*	1	*n*	$\sigma^2 + AD$
$(\beta\gamma)_{jk}$	*a*	0	0	*d*	*n*	$\sigma^2 + BCD + BC$
$(\beta\delta)_{jl}$	*a*	0	*c*	1	*n*	$\sigma^2 + BD$
$(\gamma\delta)_{kl}$	*a*	*b*	0	1	*n*	$\sigma^2 + CD$
$(\tau\beta\gamma)_{ijk}$	0	0	0	*d*	*n*	$\sigma^2 + ABCD + ABC$
$(\tau\beta\delta)_{ijl}$	0	0	*c*	1	*n*	$\sigma^2 + ABD$
$(\beta\gamma\delta)_{jkl}$	*a*	0	0	1	*n*	$\sigma^2 + BCD$
$(\tau\gamma\delta)_{ikl}$	0	*b*	0	1	*n*	$\sigma^2 + ACD$
$(\tau\beta\gamma\delta)_{ijkl}$	0	0	0	1	*n*	$\sigma^2 + ABCD$
$\varepsilon_{(ijkl)h}$	1	1	1	1	1	σ^2

There are exact tests for all effects. The results can also be generated in *Minitab* as follows:

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A         fixed      2     H     L
B         fixed      2     H     L
C         fixed      2     H     L
D        random      2     H     L

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13    1.96  0.395
B            1       0.13       0.13    0.04  0.874
C            1       1.13       1.13    0.36  0.656
D            1       0.13       0.13    0.01  0.921
A*B          1       3.13       3.13    0.11  0.795
A*C          1       3.13       3.13    1.00  0.500
A*D          1       3.13       3.13    0.25  0.622
B*C          1       3.13       3.13    1.00  0.500
B*D          1       3.13       3.13    0.25  0.622
C*D          1       3.13       3.13    0.25  0.622
A*B*C        1       3.13       3.13    1.00  0.500
A*B*D        1      28.13      28.13    2.27  0.151
A*C*D        1       3.13       3.13    0.25  0.622
B*C*D        1       3.13       3.13    0.25  0.622
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

Source      Variance Error Expected Mean Square for Each Term
           component term (using restricted model)
 1 A                  7   (16) + 8(7) + 16Q[1]
 2 B                  9   (16) + 8(9) + 16Q[2]
 3 C                 10   (16) + 8(10) + 16Q[3]
 4 D         -0.7656  16   (16) + 16(4)
 5 A*B               12   (16) + 4(12) + 8Q[5]
 6 A*C               13   (16) + 4(13) + 8Q[6]
 7 A*D       -1.1563  16   (16) + 8(7)
 8 B*C               14   (16) + 4(14) + 8Q[8]
 9 B*D       -1.1563  16   (16) + 8(9)
10 C*D       -1.1563  16   (16) + 8(10)
11 A*B*C             15   (16) + 2(15) + 4Q[11]
12 A*B*D      3.9375  16   (16) + 4(12)
13 A*C*D     -2.3125  16   (16) + 4(13)
14 B*C*D     -2.3125  16   (16) + 4(14)
15 A*B*C*D   -4.6250  16   (16) + 2(15)
16 Error     12.3750       (16)
```

13.18. Reconsider cases (c), (d) and (e) of Problem 13.17. Obtain the expected mean squares assuming the unrestricted model. You may use a computer package such as *Minitab*. Compare your results with those for the restricted model.

A is fixed and *B*, *C*, and *D* are random.

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A         fixed      2     H     L
B        random      2     H     L
C        random      2     H     L
D        random      2     H     L
```

```
Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13      **
B            1       0.13       0.13      **
C            1       1.13       1.13    0.36  0.843 x
D            1       0.13       0.13      **
A*B          1       3.13       3.13    0.11  0.796 x
A*C          1       3.13       3.13    1.00  0.667 x
A*D          1       3.13       3.13    0.11  0.796 x
B*C          1       3.13       3.13    1.00  0.667 x
B*D          1       3.13       3.13    0.11  0.796 x
C*D          1       3.13       3.13    1.00  0.667 x
A*B*C        1       3.13       3.13    1.00  0.500
A*B*D        1      28.13      28.13    9.00  0.205
A*C*D        1       3.13       3.13    1.00  0.500
B*C*D        1       3.13       3.13    1.00  0.500
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

x Not an exact F-test.

** Denominator of F-test is zero.

Source      Variance Error Expected Mean Square for Each Term
           component term (using unrestricted model)
 1 A                   *    (16) + 2(15) + 4(13) + 4(12) + 4(11) + 8(7) + 8(6)
                             + 8(5) + Q[1]
 2 B          1.3750   *    (16) + 2(15) + 4(14) + 4(12) + 4(11) + 8(9) + 8(8)
                             + 8(5) + 16(2)
 3 C         -0.1250   *    (16) + 2(15) + 4(14) + 4(13) + 4(11) + 8(10) + 8(8)
                             + 8(6) + 16(3)
 4 D          1.3750   *    (16) + 2(15) + 4(14) + 4(13) + 4(12) + 8(10) + 8(9)
                             + 8(7) + 16(4)
 5 A*B       -3.1250   *    (16) + 2(15) + 4(12) + 4(11) + 8(5)
 6 A*C        0.0000   *    (16) + 2(15) + 4(13) + 4(11) + 8(6)
 7 A*D       -3.1250   *    (16) + 2(15) + 4(13) + 4(12) + 8(7)
 8 B*C        0.0000   *    (16) + 2(15) + 4(14) + 4(11) + 8(8)
 9 B*D       -3.1250   *    (16) + 2(15) + 4(14) + 4(12) + 8(9)
10 C*D        0.0000   *    (16) + 2(15) + 4(14) + 4(13) + 8(10)
11 A*B*C      0.0000  15    (16) + 2(15) + 4(11)
12 A*B*D      6.2500  15    (16) + 2(15) + 4(12)
13 A*C*D      0.0000  15    (16) + 2(15) + 4(13)
14 B*C*D      0.0000  15    (16) + 2(15) + 4(14)
15 A*B*C*D   -4.6250  16    (16) + 2(15)
16 Error     12.3750        (16)

* Synthesized Test.

Error Terms for Synthesized Tests

Source       Error DF  Error MS  Synthesis of Error MS
 1 A             0.56         *   (5) + (6) + (7) - (11) - (12) - (13) + (15)
 2 B             0.56         *   (5) + (8) + (9) - (11) - (12) - (14) + (15)
 3 C             0.14      3.13   (6) + (8) + (10) - (11) - (13) - (14) + (15)
 4 D             0.56         *   (7) + (9) + (10) - (12) - (13) - (14) + (15)
 5 A*B           0.98     28.13   (11) + (12) - (15)
 6 A*C           0.33      3.13   (11) + (13) - (15)
 7 A*D           0.98     28.13   (12) + (13) - (15)
 8 B*C           0.33      3.13   (11) + (14) - (15)
 9 B*D           0.98     28.13   (12) + (14) - (15)
10 C*D           0.33      3.13   (13) + (14) - (15)
```

A and *B* are fixed and *C* and *D* are random.

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A         fixed      2     H     L
B         fixed      2     H     L
C        random      2     H     L
D        random      2     H     L

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1       6.13       6.13    1.96  0.604 x
B            1       0.13       0.13    0.04  0.907 x
C            1       1.13       1.13    0.36  0.843 x
D            1       0.13       0.13      **
A*B          1       3.13       3.13    0.11  0.796 x
A*C          1       3.13       3.13    1.00  0.667 x
A*D          1       3.13       3.13    0.11  0.796 x
B*C          1       3.13       3.13    1.00  0.667 x
B*D          1       3.13       3.13    0.11  0.796 x
C*D          1       3.13       3.13    1.00  0.667 x
A*B*C        1       3.13       3.13    1.00  0.500
A*B*D        1      28.13      28.13    9.00  0.205
A*C*D        1       3.13       3.13    1.00  0.500
B*C*D        1       3.13       3.13    1.00  0.500
A*B*C*D      1       3.13       3.13    0.25  0.622
Error       16     198.00      12.38
Total       31     264.88

x Not an exact F-test.

** Denominator of F-test is zero.

Source      Variance Error Expected Mean Square for Each Term
           component term (using unrestricted model)
 1 A                   *   (16) + 2(15) + 4(13) + 4(12) + 4(11) + 8(7) + 8(6)
                            + Q[1,5]
 2 B                   *   (16) + 2(15) + 4(14) + 4(12) + 4(11) + 8(9) + 8(8)
                            + Q[2,5]
 3 C         -0.1250   *   (16) + 2(15) + 4(14) + 4(13) + 4(11) + 8(10) + 8(8)
                            + 8(6) + 16(3)
 4 D          1.3750   *   (16) + 2(15) + 4(14) + 4(13) + 4(12) + 8(10) + 8(9)
                            + 8(7) + 16(4)
 5 A*B                 *   (16) + 2(15) + 4(12) + 4(11) + Q[5]
 6 A*C        0.0000   *   (16) + 2(15) + 4(13) + 4(11) + 8(6)
 7 A*D       -3.1250   *   (16) + 2(15) + 4(13) + 4(12) + 8(7)
 8 B*C        0.0000   *   (16) + 2(15) + 4(14) + 4(11) + 8(8)
 9 B*D       -3.1250   *   (16) + 2(15) + 4(14) + 4(12) + 8(9)
10 C*D        0.0000   *   (16) + 2(15) + 4(14) + 4(13) + 8(10)
11 A*B*C      0.0000  15   (16) + 2(15) + 4(11)
12 A*B*D      6.2500  15   (16) + 2(15) + 4(12)
13 A*C*D      0.0000  15   (16) + 2(15) + 4(13)
14 B*C*D      0.0000  15   (16) + 2(15) + 4(14)
15 A*B*C*D   -4.6250  16   (16) + 2(15)
16 Error     12.3750       (16)

* Synthesized Test.

Error Terms for Synthesized Tests

Source       Error DF  Error MS  Synthesis of Error MS
 1 A             0.33      3.13   (6) + (7) - (13)
 2 B             0.33      3.13   (8) + (9) - (14)
 3 C             0.14      3.13   (6) + (8) + (10) - (11) - (13) - (14) + (15)
 4 D             0.56         *   (7) + (9) + (10) - (12) - (13) - (14) + (15)
 5 A*B           0.98     28.13   (11) + (12) - (15)
 6 A*C           0.33      3.13   (11) + (13) - (15)
 7 A*D           0.98     28.13   (12) + (13) - (15)
 8 B*C           0.33      3.13   (11) + (14) - (15)
 9 B*D           0.98     28.13   (12) + (14) - (15)
10 C*D           0.33      3.13   (13) + (14) - (15)
```

(e) *A*, *B* and *C* are fixed and *D* is random.

Minitab Output

ANOVA: y versus A, B, C, D

```
Factor     Type Levels Values
A         fixed      2     H     L
B         fixed      2     H     L
C         fixed      2     H     L
D        random      2     H     L

Analysis of Variance for y

Source      DF        SS        MS       F      P
A            1      6.13      6.13    1.96  0.395
B            1      0.13      0.13    0.04  0.874
C            1      1.13      1.13    0.36  0.656
D            1      0.13      0.13      **
A*B          1      3.13      3.13    0.11  0.795
A*C          1      3.13      3.13    1.00  0.500
A*D          1      3.13      3.13    0.11  0.796 x
B*C          1      3.13      3.13    1.00  0.500
B*D          1      3.13      3.13    0.11  0.796 x
C*D          1      3.13      3.13    1.00  0.667 x
A*B*C        1      3.13      3.13    1.00  0.500
A*B*D        1     28.13     28.13    9.00  0.205
A*C*D        1      3.13      3.13    1.00  0.500
B*C*D        1      3.13      3.13    1.00  0.500
A*B*C*D      1      3.13      3.13    0.25  0.622
Error       16    198.00     12.38
Total       31    264.88

x Not an exact F-test.

** Denominator of F-test is zero.

Source      Variance Error Expected Mean Square for Each Term
           component term (using unrestricted model)
 1 A                  7   (16) + 2(15) + 4(13) + 4(12) + 8(7) + Q[1,5,6,11]
 2 B                  9   (16) + 2(15) + 4(14) + 4(12) + 8(9) + Q[2,5,8,11]
 3 C                 10   (16) + 2(15) + 4(14) + 4(13) + 8(10) + Q[3,6,8,11]
 4 D          1.3750   *   (16) + 2(15) + 4(14) + 4(13) + 4(12) + 8(10) + 8(9)
                            + 8(7) + 16(4)
 5 A*B               12   (16) + 2(15) + 4(12) + Q[5,11]
 6 A*C               13   (16) + 2(15) + 4(13) + Q[6,11]
 7 A*D       -3.1250   *   (16) + 2(15) + 4(13) + 4(12) + 8(7)
 8 B*C               14   (16) + 2(15) + 4(14) + Q[8,11]
 9 B*D       -3.1250   *   (16) + 2(15) + 4(14) + 4(12) + 8(9)
10 C*D        0.0000   *   (16) + 2(15) + 4(14) + 4(13) + 8(10)
11 A*B*C             15   (16) + 2(15) + Q[11]
12 A*B*D      6.2500  15   (16) + 2(15) + 4(12)
13 A*C*D      0.0000  15   (16) + 2(15) + 4(13)
14 B*C*D      0.0000  15   (16) + 2(15) + 4(14)
15 A*B*C*D   -4.6250  16   (16) + 2(15)
16 Error     12.3750       (16)

* Synthesized Test.

Error Terms for Synthesized Tests

Source       Error DF  Error MS  Synthesis of Error MS
 4 D             0.56         *   (7) + (9) + (10) - (12) - (13) - (14) + (15)
 7 A*D           0.98     28.13   (12) + (13) - (15)
 9 B*D           0.98     28.13   (12) + (14) - (15)
10 C*D           0.33      3.13   (13) + (14) - (15)
```

13.19. In Problem 5.19, assume that the three operators were selected at random. Analyze the data under these conditions and draw conclusions. Estimate the variance components.

Minitab Output

```
ANOVA: Score versus Cycle Time, Operator, Temperature

Factor     Type Levels Values
Cycle Ti  fixed      3    40    50    60
Operator random      3     1     2     3
Temperat  fixed      2   300   350

Analysis of Variance for Score

Source                        DF         SS         MS       F      P
Cycle Ti                       2    436.000    218.000    2.45  0.202
Operator                       2    261.333    130.667   39.86  0.000
Temperat                       1     50.074     50.074    8.89  0.096
Cycle Ti*Operator              4    355.667     88.917   27.13  0.000
Cycle Ti*Temperat              2     78.815     39.407    3.41  0.137
Operator*Temperat              2     11.259      5.630    1.72  0.194
Cycle Ti*Operator*Temperat     4     46.185     11.546    3.52  0.016
Error                         36    118.000      3.278
Total                         53   1357.333

Source                        Variance Error Expected Mean Square for Each Term
                             component term (using restricted model)
 1 Cycle Ti                              4   (8) + 6(4) + 18Q[1]
 2 Operator                     7.0772   8   (8) + 18(2)
 3 Temperat                              6   (8) + 9(6) + 27Q[3]
 4 Cycle Ti*Operator           14.2731   8   (8) + 6(4)
 5 Cycle Ti*Temperat                     7   (8) + 3(7) + 9Q[5]
 6 Operator*Temperat            0.2613   8   (8) + 9(6)
 7 Cycle Ti*Operator*Temperat   2.7562   8   (8) + 3(7)
 8 Error                        3.2778       (8)
```

The following calculations agree with the *Minitab* results:

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 3.27778$$

$$\hat{\sigma}^2_{\tau\beta\gamma} = \frac{MS_{ABC} - MS_E}{n} \qquad \hat{\sigma}^2_{\tau\beta\gamma} = \frac{11.546296 - 3.277778}{3} = 2.7562$$

$$\hat{\sigma}^2_{\tau\beta} = \frac{MS_{AB} - MS_E}{cn} \qquad \hat{\sigma}^2_{\tau\beta} = \frac{88.91667 - 3.277778}{2(3)} = 14.27315$$

$$\hat{\sigma}^2_{\beta\gamma} = \frac{MS_{BC} - MS_E}{an} \qquad \hat{\sigma}^2_{\tau\gamma} = \frac{5.629630 - 3.277778}{3(3)} = 0.26132$$

$$\hat{\sigma}^2_{\beta} = \frac{MS_B - MS_E}{acn} \qquad \hat{\sigma}^2_{\gamma} = \frac{130.66667 - 3.277778}{3(2)(3)} = 7.07716$$

13.20. Consider the three-factor factorial model

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\beta\gamma)_{jk} + \varepsilon_{ijk}$$

Assuming that all the factors are random, develop the analysis of variance table, including the expected mean squares. Propose appropriate test statistics for all effects.

Source	DF	E(MS)
A	a-1	$\sigma^2 + c\sigma^2_{\tau\beta} + bc\sigma^2_{\tau}$
B	b-1	$\sigma^2 + c\sigma^2_{\tau\beta} + a\sigma^2_{\beta\gamma} + ac\sigma^2_{\beta}$
C	c-1	$\sigma^2 + a\sigma^2_{\beta\gamma} + ab\sigma^2_{\gamma}$
AB	(a-1)(b-1)	$\sigma^2 + c\sigma^2_{\tau\beta}$
BC	(b-1)(c-1)	$\sigma^2 + a\sigma^2_{\beta\gamma}$
Error (AC + ABC)	b(a-1)(c-1)	σ^2
Total	abc-1	

There are exact tests for all effects except B. To test B, use the statistic $F = \dfrac{MS_B + MS_E}{MS_{AB} + MS_{BC}}$

13.21. The three-factor factorial model for a single replicate is

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\beta\gamma)_{jk} + (\tau\gamma)_{ik} + (\tau\beta\gamma)_{ijk} + \varepsilon_{ijk}$$

If all the factors are random, can any effects be tested? If the three-factor interaction and the $(\tau\beta)_{ij}$ interaction do not exist, can all the remaining effects be tested?

The expected mean squares are found by referring to Table 13.9, deleting the line for the error term $\varepsilon_{(ijk)l}$ and setting n=1. The three-factor interaction now cannot be tested; however, exact tests exist for the two-factor interactions and approximate F tests can be conducted for the main effects. For example, to test the main effect of A, use

$$F = \frac{MS_A + MS_{ABC}}{MS_{AB} + MS_{AC}}$$

If $(\tau\beta\gamma)_{ijk}$ and $(\tau\beta)_{ij}$ can be eliminated, the model becomes

$$y_{ijk} = \mu + \tau_i + \beta_j + \gamma_k + \left(\beta\gamma\right)_{jk} + \left(\tau\gamma\right)_{ik} + \varepsilon_{ijk}$$

For this model, the analysis of variance is

Source	DF	E(MS)
A	a-1	$\sigma^2 + b\sigma^2_{\tau\gamma} + bc\sigma^2_{\tau}$
B	b-1	$\sigma^2 + a\sigma^2_{\beta\gamma} + ac\sigma^2_{\beta}$
C	c-1	$\sigma^2 + a\sigma^2_{\beta\gamma} + b\sigma^2_{\tau\gamma} + ab\sigma^2_{\gamma}$
AC	(a-1)(c-1)	$\sigma^2 + b\sigma^2_{\tau\gamma}$
BC	(b-1)(c-1)	$\sigma^2 + a\sigma^2_{\beta\gamma}$
Error (AB + ABC)	c(a-1)(b-1)	σ^2
Total	abc-1	

There are exact tests for all effect except C. To test the main effect of C, use the statistic:

$$F = \frac{MS_C + MS_E}{MS_{BC} + MS_{AC}}$$

13.22. In Problem 5.8, assume that both machines and operators were chosen randomly. Determine the power of the test for detecting a machine effect such that $\sigma_\beta^2 = \sigma^2$, where σ_β^2 is the variance component for the machine factor. Are two replicates sufficient?

$$\lambda = \sqrt{1 + \frac{an\sigma_\beta^2}{\sigma^2 + n\sigma_{\tau\beta}^2}}$$

If $\sigma_\beta^2 = \sigma^2$, then an estimate of $\sigma^2 = \sigma_\beta^2 = 3.79$, and an estimate of $\sigma_{\tau\beta}^2 = 7.44$, from the analysis of variance table. Then

$$\lambda = \sqrt{1 + \frac{(3)(2)(3.79)}{3.79 + 2(7.44)}} = \sqrt{2.22} = 1.49$$

and the other OC curve parameters are $\upsilon_1 = 3$ and $\upsilon_2 = 6$. This results in $\beta \approx 0.75$ approximately, with $\alpha = 0.05$, or $\beta \approx 0.9$ with $\alpha = 0.01$. Two replicates does not seem sufficient.

13.23. In the two-factor mixed model analysis of variance, show that $\text{Cov}\left[(\tau\beta)_{ij}, (\tau\beta)_{i'j}\right] = -(1/a)^2_{\tau\beta\sigma}$ for $i \neq i'$.

Since $\sum_{i=1}^{a} (\tau\beta)_{ij} = 0$ (constant) we have $V\left[\sum_{i=1}^{a} (\tau\beta)_{ij}\right] = 0$, which implies that

$$\sum_{i=1}^{a} V\left(\tau\beta\right)_{ij} + 2\binom{a}{2} Cov\left[\left(\tau\beta\right)_{ij}, \left(\tau\beta\right)_{i'j}\right] = 0$$

$$a\left[\frac{a-1}{a}\right]\sigma_{\tau\beta}^2 + \frac{a!}{2!(a-2)!}(2)\,Cov\left[\left(\tau\beta\right)_{ij}, \left(\tau\beta\right)_{i'j}\right] = 0$$

$$(a-1)\sigma_{\tau\beta}^2 + a(a-1)\,Cov\left[\left(\tau\beta\right)_{ij}, \tau\left(\beta\right)_{i'j}\right] = 0$$

$$Cov\left[\tau\left(\beta\right)_{ij}, \left(\tau\beta\right)_{i'j}\right] = -\left(\frac{1}{a}\right)\sigma_{\tau\beta}^2$$

13.24. Show that the method of analysis of variance always produces unbiased point estimates of the variance component in any random or mixed model.

Let $\mathbf{g}$ be the vector of mean squares from the analysis of variance, chosen so that $E(\mathbf{g})$ does not contain any fixed effects. Let $\boldsymbol{\sigma}^2$ be the vector of variance components such that $E(\mathbf{g}) = \mathbf{A}\boldsymbol{\sigma}^2$, where $\mathbf{A}$ is a matrix of constants. Now in the analysis of variance method of variance component estimation, we equate observed and expected mean squares, i.e.

$$\mathbf{g} = \mathbf{A}\mathbf{s}^2 \Rightarrow \hat{\mathbf{s}}^2 = \mathbf{A}^{-1}\mathbf{g}$$

Since $\mathbf{A}^{-1}$ always exists then,

$$E\left(\mathbf{s}^2\right) = E\left(\mathbf{A}^{-1}\right)\mathbf{g} = \mathbf{A}^{-1}E\left(\mathbf{g}\right) = \mathbf{A}^{-1}\left(\mathbf{A}\mathbf{s}^2\right) = \mathbf{s}^2$$

Thus $\hat{\boldsymbol{\sigma}}^2$ is an unbiased estimator of $\boldsymbol{\sigma}^2$. This and other properties of the analysis of variance method are discussed by Searle (1971a).

13.25. Invoking the usual normality assumptions, find an expression for the probability that a negative estimate of a variance component will be obtained by the analysis of variance method. Using this result, write a statement giving the probability that $\hat{\sigma}_\tau^2 < 0$ in a one-factor analysis of variance. Comment on the usefulness of this probability statement.

Suppose $\hat{\sigma}^2 = \dfrac{MS_1 - MS_2}{c}$, where MS_i for i=1,2 are two mean squares and c is a constant. The probability that $\hat{\sigma}_\tau^2 < 0$ (negative) is

$$P\left\{\hat{\sigma}^2 < 0\right\} = P\left\{MS_1 - MS_2 < 0\right\} = P\left\{\frac{MS_1}{MS_2} < 1\right\} = P\left\{\frac{\frac{MS_1}{E\left(MS_1\right)}}{\frac{MS_2}{E\left(MS_2\right)}} < \frac{E\left(MS_1\right)}{E\left(MS_2\right)}\right\} = P\left\{F_{u,v} < \frac{E\left(MS_1\right)}{E\left(MS_2\right)}\right\}$$

where u is the number of degrees of freedom for MS_1 and v is the number of degrees of freedom for MS_2. For the one-way model, this equation reduces to

$$P\left\{\hat{\sigma}^2 < 0\right\} = P\left\{F_{a-1,N-a} < \frac{\sigma^2}{\sigma^2 + n\sigma_\tau^2}\right\} = P\left\{F_{a-1,N-a} < \frac{1}{1+nk}\right\}$$

where $k = \dfrac{\sigma_\tau^2}{\sigma^2}$. Using arbitrary values for some of the parameters in this equation will give an experimenter some idea of the probability of obtaining a negative estimate of $\hat{\sigma}_\tau^2 < 0$.

13.26. Analyze the data in Problem 13.1, assuming that the operators are fixed, using both the unrestricted and restricted forms of the mixed models. Compare the results obtained from the two models.

The restricted model is as follows:

Minitab Output

ANOVA: Measurement versus Part, Operator

```
Factor     Type Levels Values
Part     random     10      1     2     3     4     5     6     7
                            8     9    10
Operator  fixed      2      1     2

Analysis of Variance for Measurem

Source           DF         SS         MS       F       P
Part              9     99.017     11.002    7.33  0.000
```

```
Operator          1       0.417      0.417    0.69  0.427
Part*Operator     9       5.417      0.602    0.40  0.927
Error            40      60.000      1.500
Total            59     164.850

Source           Variance Error Expected Mean Square for Each Term
                component term (using restricted model)
 1 Part            1.5836   4   (4) + 6(1)
 2 Operator                 3   (4) + 3(3) + 30Q[2]
 3 Part*Operator  -0.2994   4   (4) + 3(3)
 4 Error           1.5000       (4)
```

The second approach is the unrestricted mixed model.

Minitab Output

```
ANOVA: Measurement versus Part, Operator

Factor     Type Levels Values
Part     random     10     1     2     3     4     5     6     7
                           8     9    10
Operator  fixed      2     1     2

Analysis of Variance for Measurem

Source           DF         SS         MS       F      P
Part              9     99.017     11.002   18.28  0.000
Operator          1      0.417      0.417    0.69  0.427
Part*Operator     9      5.417      0.602    0.40  0.927
Error            40     60.000      1.500
Total            59    164.850

Source           Variance Error Expected Mean Square for Each Term
                component term (using unrestricted model)
 1 Part            1.7333   3   (4) + 3(3) + 6(1)
 2 Operator                 3   (4) + 3(3) + Q[2]
 3 Part*Operator  -0.2994   4   (4) + 3(3)
 4 Error           1.5000       (4)
```

Source	Sum of Squares	DF	Mean Square	E(MS)	*F*-test	*F*
A	99.016667	*a*-1=9	11.00185	$\sigma^2 + n\sigma_{\tau\beta}^2 + bn\sigma_\tau^2$	$F = \frac{MS_A}{MS_{AB}}$	18.28
B	0.416667	*b*-1=1	0.416667	$\sigma^2 + n\sigma_{\tau\beta}^2 + an\frac{\sum_{i=1}^{b}\beta_i^2}{b-1}$	$F = \frac{MS_B}{MS_{AB}}$	0.692
AB	5.416667	(*a*-1)(*b*-1)=9	0.60185	$\sigma^2 + n\sigma_{\tau\beta}^2$	$F = \frac{MS_{AB}}{MS_E}$	0.401
Error	60.000000	40	1.50000	σ^2		
Total	164.85000	*nabc*-1=59				

In the unrestricted model, the *F*-test for *A* is different. The *F*-test for *A* in the unrestricted model should generally be more conservative, since MS_{AB} will generally be larger than MS_E. However, this is not the case with this particular experiment.

13.27. Consider the two-factor mixed model. Show that the standard error of the fixed factor mean (e.g. *A*) is $\left[MS_{AB}/bn\right]^{1/2}$.

The standard error is often used in Duncan's Multiple Range test. Duncan's Multiple Range Test requires the variance of the difference in two means, say

$$V\left(\bar{y}_{i..} - \bar{y}_{m..}\right)$$

where rows are fixed and columns are random. Now, assuming all model parameters to be independent, we have the following:

$$\left(\bar{y}_{i..} - \bar{y}_{m..}\right) = \tau_i - \tau_m + \frac{1}{b}\sum_{j=1}^{b}(\tau\beta)_{ij} - \frac{1}{b}\sum_{j=1}^{b}(\tau\beta)_{mj} + \frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n}\varepsilon_{ijk} - \frac{1}{bn}\sum_{j=1}^{b}\sum_{k=1}^{n}\varepsilon_{mjk}$$

and

$$V\left(\bar{y}_{i..} - \bar{y}_{m..}\right) = \left(\frac{1}{b}\right)^2 b\sigma_{\tau\beta}^2 + \left(\frac{1}{b}\right)^2 b\sigma_{\tau\beta}^2 + \left(\frac{1}{bn}\right)^2 bn\sigma^2 + \left(\frac{1}{bn}\right)^2 bn\sigma^2 = \frac{2\left(\sigma^2 + n\sigma_{\tau\beta}^2\right)}{bn}$$

Since MS_{AB} estimates $\sigma^2 + n\sigma_{\tau\beta}^2$, we would use

$$\frac{2MS_{AB}}{bn}$$

as the standard error to test the difference. However, the table of ranges for Duncan's Multiple Range test already includes the constant 2.

13.28. Consider the variance components in the random model from Problem 13.1.

(a) Find an exact 95 percent confidence interval on σ^2.

$$\frac{f_E MS_E}{\chi^2_{\alpha/2, f_E}} \le \sigma^2 \le \frac{f_E MS_E}{\chi^2_{1-\alpha/2, f_E}}$$

$$\frac{(40)(1.5)}{59.34} \le \sigma^2 \le \frac{(40)(1.5)}{24.43}$$

$$1.011 \le \sigma^2 \le 2.456$$

(b) Find approximate 95 percent confidence intervals on the other variance components using the Satterthwaite method.

$\hat{\sigma}_{\tau\beta}^2$ and $\hat{\sigma}_{\tau}^2$ are negative, and the Satterthwaithe method does not apply. The confidence interval on $\hat{\sigma}_{\beta}^2$ is

$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_{AB}}{an} \qquad \hat{\sigma}_\beta^2 = \frac{11.001852 - 0.6018519}{2(3)} = 1.7333$$

$$r = \frac{\left(MS_B - MS_{AB}\right)^2}{\frac{MS_B^2}{(b-1)} + \frac{MS_{AB}^2}{(a-1)(b-1)}} = \frac{\left(11.001852 - 0.6018519\right)^2}{\frac{11.001852^2}{(9)} + \frac{0.6018519^2}{(1)(9)}} = 8.01826$$

$$\frac{r\hat{\sigma}_O^2}{\chi^2_{\alpha/2,r}} \le \sigma_\beta^2 \le \frac{r\hat{\sigma}_\beta^2}{\chi^2_{1-\alpha/2,r}}$$

$$\frac{(8.01826)(1.7333)}{17.55752} \le \sigma_\beta^2 \le \frac{(8.01826)(1.7333)}{2.18950}$$

$$0.79157 \le \sigma_\beta^2 \le 6.34759$$

13.29. Use the experiment described in Problem 5.8 and assume that both factors are random. Find an exact 95 percent confidence interval on σ^2. Construct approximate 95 percent confidence interval on the other variance components using the Satterthwaite method.

$$\hat{\sigma}^2 = MS_E \qquad \hat{\sigma}^2 = 3.79167$$

$$\frac{f_E MS_E}{\chi^2_{\alpha/2,f_E}} \le \sigma^2 \le \frac{f_E MS_E}{\chi^2_{1-\alpha/2,f_E}}$$

$$\frac{(12)(3.79167)}{23.34} \le \sigma^2 \le \frac{(12)(3.79167)}{4.40}$$

$$1.9494 \le \sigma^2 \le 10.3409$$

Satterthwaite Method:

$$\hat{\sigma}_{\tau\beta}^2 = \frac{MS_{AB} - MS_E}{n} \qquad \hat{\sigma}_{\tau\beta}^2 = \frac{7.44444 - 3.79167}{2} = 1.82639$$

$$r = \frac{(MS_{AB} - MS_E)^2}{\frac{MS_{AB}^2}{(a-1)(b-1)} + \frac{MS_E^2}{df_E}} = \frac{(7.44444 - 3.79167)^2}{\frac{7.44444^2}{(2)(3)} + \frac{3.79167^2}{(12)}} = 1.27869$$

$$\frac{r\hat{\sigma}_\beta^2}{\chi^2_{\alpha/2,r}} \le \sigma_\beta^2 \le \frac{r\hat{\sigma}_\beta^2}{\chi^2_{1-\alpha/2,r}}$$

$\chi^2_{\alpha/2,r}$ and $\chi^2_{1-\alpha/2,r}$ were estimated by extrapolating 1.27869 between degrees of freedom of one and two in *Microsoft Excel.* More precise methods can be used as well.

$$\frac{(1.27869)(1.82639)}{5.67991} \le \sigma_\beta^2 \le \frac{(1.27869)(1.82639)}{0.014820}$$

$$0.41117 \le \sigma_\beta^2 \le 158.58172$$

$\hat{\sigma}_\beta^2 < 0$, this variance component does not have a confidence interval using Satterthwaite's Method.

$$\hat{\sigma}_\tau^2 = \frac{MS_A - MS_{AB}}{bn} \qquad \hat{\sigma}_\tau^2 = \frac{80.16667 - 7.44444}{4(2)} = 9.09028$$

$$r = \frac{\left(MS_A - MS_{AB}\right)^2}{\frac{MS_A^2}{(a-1)} + \frac{MS_{AB}^2}{(a-1)(b-1)}} = \frac{\left(80.16667 - 7.44444\right)^2}{\frac{80.16667^2}{(2)} + \frac{7.44444^2}{(2)(3)}} = 1.64108$$

$$\frac{r\hat{\sigma}_\tau^2}{\chi^2_{\alpha/2,r}} \le \sigma_\tau^2 \le \frac{r\hat{\sigma}_\tau^2}{\chi^2_{1-\alpha/2,r}}$$

$$\frac{(1.64108)(9.09028)}{6.53293} \le \sigma_\tau^2 \le \frac{(1.64108)(9.09028)}{0.03281}$$

$\chi^2_{\alpha/2,r}$ and $\chi^2_{1-\alpha/2,r}$ were estimated by extrapolating 1.64108 between degrees of freedom of one and two in *Microsoft Excel.* More precise methods can be used as well.

$$2.28349 \le \sigma_\tau^2 \le 454.61891$$

13.30. Consider the three-factor experiment in Problem 5.19 and assume that operators were selected at random. Find an approximate 95 percent confidence interval on the operator variance component.

$$\hat{\sigma}_\beta^2 = \frac{MS_B - MS_E}{acn} \qquad \hat{\sigma}_\beta^2 = \frac{130.66667 - 3.277778}{2(3)(3)} = 7.07716$$

$$r = \frac{\left(MS_B - MS_E\right)^2}{\frac{MS_B^2}{(b-1)} + \frac{MS_E^2}{df_E}} = \frac{\left(130.66667 - 3.27778\right)^2}{\frac{130.66667^2}{(2)} + \frac{3.27778^2}{(36)}} = 1.90085$$

$$\frac{r\hat{\sigma}_\beta^2}{\chi^2_{\alpha/2,r}} \le \sigma_\beta^2 \le \frac{r\hat{\sigma}_\beta^2}{\chi^2_{1-\alpha/2,r}}$$

$\chi^2_{\alpha/2,r}$ and $\chi^2_{1-\alpha/2,r}$ were estimated by extrapolating 1.90085 between degrees of freedom of one and two in *Microsoft Excel.* More precise methods can be used as well.

$$\frac{(1.90085)(7.07716)}{7.14439} \le \sigma_\beta^2 \le \frac{(1.90085)(7.07716)}{0.04571}$$

$$1.88296 \le \sigma_\beta^2 \le 294.28720$$

13.31. Rework Problem 13.28 using the modified large-sample approach described in Section 13.7.2. Compare the two sets of confidence intervals obtained and discuss.

$$\hat{\sigma}_O^2 = \hat{\sigma}_\beta^2 = \frac{MS_B - MS_{AB}}{an} \qquad \hat{\sigma}_O^2 = \frac{11.001852 - 0.6018519}{2(3)} = 1.7333$$

$$G_1 = 1 - \frac{1}{F_{0.05,9,\infty}} = 1 - \frac{1}{1.88} = 0.46809$$

$$H_1 = \frac{1}{F_{.95,9_i,\infty}} - 1 = \frac{1}{\frac{\chi^2_{.95,9}}{9}} - 1 = \frac{1}{0.370} - 1 = 1.7027$$

$$G_{ij} = \frac{\left(F_{\alpha,f_i,f_j} - 1\right)^2 - G_1^2 F_{\alpha,f_i,f_j} - H_1^2}{F_{\alpha,f_i,f_j}} = \frac{(3.18-1)^2 - (0.46809)^2(3.18) - 1.7027^2}{3.18} = 0.36366$$

$$V_L = G_1^2 c_1^2 MS_B^2 + H_1^2 c_2^2 MS_{AB}^2 + G_{11} c_1 c_2 MS_B MS_{AB}$$

$$V_L = (0.46809)^2 \left(\frac{1}{6}\right)^2 (11.00185)^2 + (1.7027)^2 \left(\frac{1}{6}\right)^2 (0.60185)^2 + (0.36366)\left(\frac{1}{6}\right)\left(\frac{1}{6}\right)(11.00185)(0.60185)$$

$$V_L = 0.83275$$

$$L = \hat{\sigma}_\beta^2 - \sqrt{V_L} = 1.7333 - \sqrt{0.83275} = 0.82075$$

13.32. Rework Problem 13.28 using the modified large-sample method described in Section 13.7.2. Compare this confidence interval with the one obtained previously and discuss.

$$\hat{\sigma}_\gamma^2 = \frac{MS_C - MS_E}{abn} \qquad \hat{\sigma}_\gamma^2 = \frac{130.66667 - 3.277778}{2(3)(3)} = 7.07716$$

$$G_1 = 1 - \frac{1}{F_{0.05,3,\infty}} = 1 - \frac{1}{2.60} = 0.61538$$

$$H_1 = \frac{1}{F_{.95,36,\infty}} - 1 = \frac{1}{\frac{\chi^2_{.95,36}}{36}} - 1 = .\frac{1}{0.64728} - 1 = 0.54493$$

$$G_{ij} = \frac{\left(F_{\alpha,f_i,f_j} - 1\right)^2 - G_1^2 F_{\alpha,f_i,f_j} - H_1^2}{F_{\alpha,f_i,f_j}} = \frac{(2.88-1)^2 - (0.61538)^2(2.88) - 0.54493^2}{2.88} = 0.74542$$

$$V_L = G_1^2 c_1^2 MS_B^2 + H_1^2 c_2^2 MS_{AB}^2 + G_{11} c_1 c_2 MS_B MS_{AB}$$

$$V_L = (0.61538)^2 \left(\frac{1}{18}\right)^2 (130.66667)^2 + (0.54493)^2 \left(\frac{1}{18}\right)^2 (3.27778)^2 + (0.74542)\left(\frac{1}{18}\right)\left(\frac{1}{18}\right)(130.66667)(3.27$$

$$V_L = 20.95112$$

$$L = \hat{\sigma}_\gamma^2 - \sqrt{V_L} = 7.07716 - \sqrt{20.95112} = 2.49992$$

13.33 Consider the experiment described in Problem 5.8. Estimate the variance component using the REML method. Compare the CIs to the approximate CIs found in Problem 13.29.

The JMP REML analysis below was performed with both factors, Operator and Machine, as random.

The CIs for the error variance are similar to those found in Problem 13.29. The upper CIs for the other variance components are much larger than those estimated in the JMP REML output below.

JMP Output

RSquare	0.742044
RSquare Adj	0.742044
Root Mean Square Error	1.94722
Mean of Response	112.2917
Observations (or Sum Wgts)	24

Parameter Estimates

Term	**Estimate**	**Std Error**	**DFDen**	**t Ratio**	**Prob>|t|**
Intercept	112.29167	1.789728	1.831	62.74	0.0005*

REML Variance Component Estimates

Random Effect	**Var Ratio**	**Var Component**	**Std Error**	**95% Lower**	**95% Upper**	**Pct of Total**
Operator	2.3974359	9.0902778	10.035225	-10.5784	28.758958	64.198
Machine	-0.144689	-0.548611	0.9124188	-2.336919	1.239697	-3.874
Operator*Machine	0.481685	1.8263889	2.2841505	-2.650464	6.3032416	12.898
Residual		3.7916667	1.5479414	1.9497217	10.332013	26.778
Total		14.159722				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	**Operator**	**Machine**	**Operator*Machine**	**Residual**
Operator	100.70575	0.3848594	-1.154578	2.151e-12
Machine	0.3848594	0.8325081	-1.539438	-1.17e-13
Operator*Machine	-1.154578	-1.539438	5.2173434	-1.198061
Residual	2.151e-12	-1.17e-13	-1.198061	2.3961227

13.34 Consider the experiment described in Problem 13.1. Analyze the data using REML. Compare the CIs to those obtained in Problem 13.28.

The JMP REML analysis below was performed with both factors, Part Number and Operator, as random.

The CIs for the Operator and Part Number Operator interaction were not calculated in Problem 13.28 due to negative estimates for the corresponding variance components. The error variance estimates and CIs found in the JMP REML output below are the same as those calculated in Problem 13.28. The upper CI for the Part Number variance estimated with the Satterthwaite method in Problem 13.28 is approximately twice the value estimated in the JMP REML analysis.

JMP Output

RSquare	0.388009
RSquare Adj	0.388009
Root Mean Square Error	1.224745
Mean of Response	49.95
Observations (or Sum Wgts)	60

Parameter Estimates

Term	**Estimate**	**Std Error**	**DFDen**	**t Ratio**	**Prob>|t|**
Intercept	49.95	0.424591	8.563	117.64	<.0001*

REML Variance Component Estimates

Random Effect	**Var Ratio**	**Var Component**	**Std Error**	**95% Lower**	**95% Upper**	**Pct of Total**
Part Number	1.1555556	1.7333333	0.8656795	0.0366326	3.430034	59.203
Operator	-0.004115	-0.006173	0.0218	-0.0489	0.0365544	-0.211
Part Number*Operator	-0.199588	-0.299383	0.1464372	-0.586394	-0.012371	-10.226
Residual		1.5	0.3354102	1.0110933	2.4556912	51.233
Total		2.9277778				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	Part Number	Operator	Part Number*Operator	Residual
Part Number	0.749401	0.0004472	-0.004472	9.256e-14
Operator	0.0004472	0.0004752	-0.000894	-3.32e-16
Part Number*Operator	-0.004472	-0.000894	0.0214438	-0.0375
Residual	9.256e-14	-3.32e-16	-0.0375	0.1125

13.35 Rework Problem 13.31 using REML. Compare all sets of CIs for the variance components.

JMP uses the unrestricted approach for estimating variance components with mixed models. The JMP REML output is shown below.

The lower confidence interval on σ_β^2 (Operator) comparisons between the Satterthwaite method in Problem 13.28, the modified large sample approach in Problem 13.31, and the JMP REML output below are:

Satterthwaite $0.79157 \le \sigma_\beta^2 \le 6.34759$

Modified $0.82075 \le \sigma_\beta^2$

REML $0.03663 \le \sigma_\beta^2 \le 3.43003$

JMP Output

RSquare	0.388009
RSquare Adj	0.388009
Root Mean Square Error	1.224745
Mean of Response	49.95
Observations (or Sum Wgts)	60

Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|
Intercept	49.95	0.424591	8.563	117.64	<.0001*

REML Variance Component Estimates

Random Effect	Var Ratio	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Part Number	1.1555556	1.7333333	0.8656795	0.0366326	3.430034	59.203
Operator	-0.004115	-0.006173	0.0218	-0.0489	0.0365544	-0.211
Part Number*Operator	-0.199588	-0.299383	0.1464372	-0.586394	-0.012371	-10.226
Residual		1.5	0.3354102	1.0110933	2.4556912	51.233
Total		2.9277778				100.000

Covariance Matrix of Variance Component Estimates

Random Effect	Part Number	Operator	Part Number*Operator	Residual
Part Number	0.749401	0.0004472	-0.004472	9.256e-14
Operator	0.0004472	0.0004752	-0.000894	-3.32e-16
Part Number*Operator	-0.004472	-0.000894	0.0214438	-0.0375
Residual	9.256e-14	-3.32e-16	-0.0375	0.1125

Chapter 14
Nested and Split-Plot Designs
Solutions

In this chapter we have not shown residual plots and other diagnostics to conserve space. A complete analysis would, of course, include these model adequacy checking procedures.

14.1. A rocket propellant manufacturer is studying the burning rate of propellant from three production processes. Four batches of propellant are randomly selected from the output of each process and three determinations of burning rate are made on each batch. The results follow. Analyze the data and draw conclusions.

	Process 1				Process 2				Process 3			
Batch	1	2	3	4	1	2	3	4	1	2	3	4
	25	19	15	15	19	23	18	35	14	35	38	25
	30	28	17	16	17	24	21	27	15	21	54	29
	26	20	14	13	14	21	17	25	20	24	50	33

Minitab Output

ANOVA: Burn Rate versus Process, Batch

```
Factor           Type Levels Values
Process         fixed      3     1     2     3
Batch(Process) random      4     1     2     3     4

Analysis of Variance for Burn Rat

Source            DF         SS         MS       F      P
Process            2     676.06     338.03    1.46  0.281
Batch(Process)     9    2077.58     230.84   12.20  0.000
Error             24     454.00      18.92
Total             35    3207.64

Source            Variance Error Expected Mean Square for Each Term
                 component term (using restricted model)
 1 Process                   2   (3) + 3(2) + 12Q[1]
 2 Batch(Process)    70.64   3   (3) + 3(2)
 3 Error             18.92       (3)
```

There is no significant effect on mean burning rate among the different processes; however, different batches from the same process have significantly different burning rates.

14.2. The surface finish of metal parts made on four machines is being studied. An experiment is conducted in which each machine is run by three different operators and two specimens from each operator are collected and tested. Because of the location of the machines, different operators are used on each machine, and the operators are chosen at random. The data are shown in the following table. Analyze the data and draw conclusions.

	Machine 1			Machine 2			Machine 3			Machine 4		
Operator	1	2	3	1	2	3	1	2	3	1	2	3
	79	94	46	92	85	76	88	53	46	36	40	62
	62	74	57	99	79	68	75	56	57	53	56	47

Minitab Output

```
ANOVA: Finish versus Machine, Operator

Factor              Type Levels Values
Machine            fixed      4     1     2     3     4
Operator(Machine) random      3     1     2     3

Analysis of Variance for Finish

Source              DF         SS         MS       F      P
Machine              3    3617.67    1205.89    3.42  0.073
Operator(Machine)    8    2817.67     352.21    4.17  0.013
Error               12    1014.00      84.50
Total               23    7449.33

Source              Variance Error Expected Mean Square for Each Term
                   component term (using restricted model)
 1 Machine                     2   (3) + 2(2) + 6Q[1]
 2 Operator(Machine)   133.85   3   (3) + 2(2)
 3 Error                84.50       (3)
```

There is a slight effect on surface finish due to the different machines; however, the different operators running the same machine have significantly different surface finish.

14.3. A manufacturing engineer is studying the dimensional variability of a particular component that is produced on three machines. Each machine has two spindles, and four components are randomly selected from each spindle. The results follow. Analyze the data, assuming that machines and spindles are fixed factors.

	Machine 1		Machine 2		Machine 3	
Spindle	1	2	1	2	1	2
	12	8	14	12	14	16
	9	9	15	10	10	15
	11	10	13	11	12	15
	12	8	14	13	11	14

Minitab Output

```
ANOVA: Variability versus Machine, Spindle

Factor             Type Levels Values
Machine           fixed      3     1     2     3
Spindle(Machine)  fixed      2     1     2

Analysis of Variance for Variabil

Source             DF         SS         MS       F      P
Machine             2     55.750     27.875   18.93  0.000
Spindle(Machine)    3     43.750     14.583    9.91  0.000
Error              18     26.500      1.472
Total              23    126.000
```

There is a significant effect on dimensional variability due to the machine and spindle factors.

14.4. To simplify production scheduling, an industrial engineer is studying the possibility of assigning one time standard to a particular class of jobs, believing that differences between jobs is negligible. To see if this simplification is possible, six jobs are randomly selected. Each job is given to a different group of three operators. Each operator completes the job twice at different times during the week, and the following results are obtained. What are your conclusions about the use of a common time standard for all jobs in this class? What value would you use for the standard?

	Operator 1		Operator 2		Operator 3	
Job	1	2	1	2	1	2
1	158.3	159.4	159.2	159.6	158.9	157.8
2	154.6	154.9	157.7	156.8	154.8	156.3
3	162.5	162.6	161.0	158.9	160.5	159.5
4	160.0	158.7	157.5	158.9	161.1	158.5
5	156.3	158.1	158.3	156.9	157.7	156.9
6	163.7	161.0	162.3	160.3	162.6	161.8

Minitab Output

ANOVA: Time versus Job, Operator

```
Factor          Type Levels Values
Job           random      6     1     2     3     4     5     6
Operator(Job) random      3     1     2     3

Analysis of Variance for Time

Source           DF        SS        MS       F      P
Job               5   148.111    29.622   17.21  0.000
Operator(Job)    12    20.653     1.721    1.58  0.186
Error            18    19.665     1.092
Total            35   188.430

Source           Variance Error Expected Mean Square for Each Term
                component term (using restricted model)
 1 Job             4.6502   2   (3) + 2(2) + 6(1)
 2 Operator(Job)   0.3143   3   (3) + 2(2)
 3 Error           1.0925       (3)
```

The jobs differ significantly; the use of a common time standard would likely not be a good idea.

14.5. Consider the three-stage nested design shown in Figure 14.5 to investigate alloy hardness. Using the data that follow, analyze the design, assuming that alloy chemistry and heats are fixed factors and ingots are random. Use the restricted form of the mixed model.

	Alloy Chemistry											
	1						2					
Heats	1		2		3		1		2		3	
Ingots	1	2	1	2	1	2	1	2	1	2	1	2
	40	27	95	69	65	78	22	23	83	75	61	35
	63	30	67	47	54	45	10	39	62	64	77	42

Minitab Output

ANOVA: Hardness versus Alloy, Heat, Ingot

```
Factor              Type Levels Values
Alloy              fixed      2     1     2
Heat(Alloy)        fixed      3     1     2     3
Ingot(Alloy Heat) random      2     1     2

Analysis of Variance for Hardness

Source              DF        SS        MS       F      P
Alloy                1     315.4     315.4    0.85  0.392
Heat(Alloy)          4    6453.8    1613.5    4.35  0.055
Ingot(Alloy Heat)    6    2226.3     371.0    2.08  0.132
Error               12    2141.5     178.5
Total               23   11137.0
```

```
Source               Variance Error Expected Mean Square for Each Term
                     component term (using restricted model)
 1 Alloy                        3   (4) + 2(3) + 12Q[1]
 2 Heat(Alloy)                  3   (4) + 2(3) + 4Q[2]
 3 Ingot(Alloy Heat)    96.29   4   (4) + 2(3)
 4 Error               178.46       (4)
```

Alloy hardness differs significantly due to the different heats within each alloy.

14.6. Reanalyze the experiment in Problem 14.5 using the unrestricted form of the mixed model. Comment on any differences you observe between the restricted and unrestricted model results. You may use a computer software package.

Minitab Output

ANOVA: Hardness versus Alloy, Heat, Ingot

```
Factor              Type Levels Values
Alloy              fixed      2     1     2
Heat(Alloy)        fixed      3     1     2     3
Ingot(Alloy Heat) random      2     1     2

Analysis of Variance for Hardness

Source               DF         SS         MS       F      P
Alloy                 1      315.4      315.4    0.85  0.392
Heat(Alloy)           4     6453.8     1613.5    4.35  0.055
Ingot(Alloy Heat)     6     2226.3      371.0    2.08  0.132
Error                12     2141.5      178.5
Total                23    11137.0

Source               Variance Error Expected Mean Square for Each Term
                     component term (using unrestricted model)
 1 Alloy                        3   (4) + 2(3) + Q[1,2]
 2 Heat(Alloy)                  3   (4) + 2(3) + Q[2]
 3 Ingot(Alloy Heat)    96.29   4   (4) + 2(3)
 4 Error               178.46       (4)
```

14.7. Derive the expected means squares for a balanced three-stage nested design, assuming that *A* is fixed and that *B* and *C* are random. Obtain formulas for estimating the variance components.

The expected mean squares can be generated in Minitab as follows:

Minitab Output

ANOVA: y versus A, B, C

```
Factor     Type Levels Values
A         fixed      2    -1     1
B(A)     random      2    -1     1
C(A B)   random      2    -1     1

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1      0.250      0.250    0.06  0.831
B(A)         2      8.500      4.250    0.35  0.726
C(A B)       4     49.000     12.250    2.13  0.168
Error        8     46.000      5.750
Total       15    103.750

Source      Variance Error Expected Mean Square for Each Term
           component term (using restricted model)
 1 A                   2   (4) + 2(3) + 4(2) + 8Q[1]
 2 B(A)       -2.000   3   (4) + 2(3) + 4(2)
 3 C(A B)      3.250   4   (4) + 2(3)
 4 Error       5.750       (4)
```

14.8. Repeat Problem 14.7 assuming the unrestricted form of the mixed model. You may use a computer software package. Comment on any differences you observe between the restricted and unrestricted model analysis and conclusions.

Minitab Output

ANOVA: y versus A, B, C

```
Factor     Type Levels Values
A         fixed      2    -1     1
B(A)     random      2    -1     1
C(A B)   random      2    -1     1

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1      0.250      0.250    0.06  0.831
B(A)         2      8.500      4.250    0.35  0.726
C(A B)       4     49.000     12.250    2.13  0.168
Error        8     46.000      5.750
Total       15    103.750

Source      Variance Error Expected Mean Square for Each Term
           component term (using unrestricted model)
 1 A                  2   (4) + 2(3) + 4(2) + Q[1]
 2 B(A)       -2.000   3   (4) + 2(3) + 4(2)
 3 C(A B)      3.250   4   (4) + 2(3)
 4 Error       5.750       (4)
```

In this case there is no difference in results between the restricted and unrestricted models.

14.9. Derive the expected means squares for a balanced three-stage nested design if all three factors are random. Obtain formulas for estimating the variance components. Assume the restricted form of the mixed model.

The expected mean squares can be generated in Minitab as follows:

Minitab Output

ANOVA: y versus A, B, C

```
Factor     Type Levels Values
A        random      2    -1     1
B(A)     random      2    -1     1
C(A B)   random      2    -1     1

Analysis of Variance for y

Source      DF         SS         MS       F      P
A            1      0.250      0.250    0.06  0.831
B(A)         2      8.500      4.250    0.35  0.726
C(A B)       4     49.000     12.250    2.13  0.168
Error        8     46.000      5.750
Total       15    103.750

Source      Variance Error Expected Mean Square for Each Term
           component term (using unrestricted model)
 1 A         -0.5000   2   (4) + 2(3) + 4(2) + 8(1)
 2 B(A)      -2.0000   3   (4) + 2(3) + 4(2)
 3 C(A B)     3.2500   4   (4) + 2(3)
 4 Error      5.7500       (4)
```

14.10. Verify the expected mean squares given in Table 14.1.

Factor	F a i	F b j	R n l	E(MS)
τ_i	0	b	n	$\sigma^2 + \frac{bn}{a-1}\sum \tau_i^2$
$\beta_{j(i)}$	1	0	n	$\sigma^2 + \frac{n}{a(b-1)}\sum\sum \beta_{j(i)}^2$
$\varepsilon_{(ijk)l}$	1	1	1	σ^2

Factor	R a i	R b j	R n l	E(MS)
τ_i	1	b	n	$\sigma^2 + n\sigma_\beta^2 + bn\sigma_\tau^2$
$\beta_{j(i)}$	1	1	n	$\sigma^2 + n\sigma_\beta^2$
$\varepsilon_{(ijk)l}$	1	1	1	σ^2

Factor	F a i	R b j	R n l	E(MS)
τ_i	0	b	n	$\sigma^2 + n\sigma_\beta^2 + \frac{bn}{a-1}\sum \tau_i^2$
$\beta_{j(i)}$	1	1	n	$\sigma^2 + n\sigma_\beta^2$
$\varepsilon_{(ijk)l}$	1	1	1	σ^2

14.11. ***Unbalanced nested designs.*** Consider an unbalanced two-stage nested design with b_j levels of B under the ith level of A and n_{ij} replicates in the ijth cell.

(a) Write down the least squares normal equations for this situation. Solve the normal equations.

The least squares normal equations are:

$$\mu = n_{..}\hat{\mu} + \sum_{i=1}^{a} n_{i.}\hat{\tau}_i + \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}\hat{\beta}_{j(i)} = y_{...}$$

$$\tau_i = n_{i.}\hat{\mu} + n_{i.}\hat{\tau}_i + \sum_{j=1}^{b_i} n_{ij}\hat{\beta}_{j(i)} = y_{i..}\text{ , for } i = 1,2,\ldots,a$$

$$\beta_{j(i)} = n_{ij}\hat{\mu} + n_{ij}\hat{\tau}_i + n_{ij}\hat{\beta}_{j(i)} = y_{ij.}\text{, for } i = 1,2,\ldots,a \text{ and } j = 1,2,\ldots,b_i$$

There are 1+a+b equations in 1+a+b unknowns. However, there are a+1 linear dependencies in these equations, and consequently, a+1 side conditions are needed to solve them. Any convenient set of a+1 linearly independent equations can be used. The easiest set is $\hat{\mu} = 0$, $\hat{\tau}_i = 0$, for i=1,2,…,a. Using these conditions we get

$$\hat{\mu} = 0\text{ , } \hat{\tau}_i = 0\text{ , } \hat{\beta}_{j(i)} = \bar{y}_{ij.}$$

as the solution to the normal equations. See Searle (1971) for a full discussion.

(b) Construct the analysis of variance table for the unbalanced two-stage nested design.

The analysis of variance table is

Source	SS	DF
A	$\sum_{i=1}^{a} \frac{y_{i..}^2}{n_{i.}} - \frac{y_{...}^2}{n_{..}}$	a-1
B	$\sum_{i=1}^{a} \sum_{j=1}^{b_i} \frac{y_{ij.}^2}{n_{ij}} - \sum_{i-1}^{a} \frac{y_{i..}^2}{n_{i.}}$	$b.$-a
Error	$\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} y_{ijk}^2 - \sum_{i=1}^{a} \sum_{j=1}^{b_i} \frac{y_{ij.}^2}{n_{ij}}$	$n..$-b
Total	$\sum_{i=1}^{a} \sum_{j=1}^{b_i} \sum_{k=1}^{n_{ij}} y_{ijk}^2 - \frac{y_{...}^2}{n_{..}}$	$n..$-1

(c) Analyze the following data, using the results in part (b).

Factor *A*	1		2		
Factor *B*	1	2	1	2	3
	6	3	5	2	1
	4	1	7	4	0
	8		9	3	-3
			6		

Note that a=2, b_1=2, b_2=3, $b.$=b_1+b_2=5, n_{11}=3, n_{12}=2, n_{21}=4, n_{22}=3 and n_{23}=3

Source	SS	DF	MS
A	0.13	1	0.13
B	153.78	3	51.26
Error	35.42	10	3.54
Total	189.33	14	

The analysis can also be performed in Minitab as follows. The adjusted sum of squares is utilized by Minitab's general linear model routine.

Minitab Output

General Linear Model: y versus A, B

```
Factor     Type Levels Values
A         fixed      2 1 2
B(A)      fixed      5 1 2 1 2 3

Analysis of Variance for y, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F       P
A           1      0.133      0.898      0.898    0.25   0.625
B(A)        3    153.783    153.783     51.261   14.47   0.001
Error      10     35.417     35.417      3.542
Total      14    189.333
```

14.12. ***Variance components in the unbalanced two-stage nested design.*** Consider the model

$$y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \varepsilon_{k(ij)} \quad \begin{cases} i = 1,2,\ldots,a \\ j = 1,2,\ldots,b_i \\ k = 1,2,\ldots,n_{ij} \end{cases}$$

where *A* and *B* are random factors. Show that

$$E(MS_A) = \sigma^2 + c_1\sigma_\beta^2 + c_2\sigma_\tau^2$$
$$E(MS_{B(A)}) = \sigma^2 + c_0\sigma_\beta^2$$
$$E(MS_E) = \sigma^2$$

where

$$c_0 = \frac{N - \sum_{i=1}^{a}\left(\sum_{j=1}^{b_i}\frac{n_{ij}^2}{n_{i.}}\right)}{b-a}$$

$$c_1 = \frac{\sum_{i=1}^{a}\left(\sum_{j=1}^{b_i}\frac{n_{ij}^2}{n_{i.}}\right) - \sum_{i=1}^{a}\sum_{j=1}^{b_i}\frac{n_{ij}^2}{N}}{a-1}$$

$$c_2 = \frac{N - \frac{\sum_{i=1}^{a} n_{i.}^2}{N}}{a-1}$$

See "Variance Component Estimation in the 2-way Nested Classification," by S.R. Searle, *Annals of Mathematical Statistics*, Vol. 32, pp. 1161-1166, 1961. A good discussion of variance component estimation from unbalanced data is in Searle (1971a).

14.13. A process engineer is testing the yield of a product manufactured on three machines. Each machine can be operated at two power settings. Furthermore, a machine has three stations on which the product is formed. An experiment is conducted in which each machine is tested at both power settings, and three observations on yield are taken from each station. The runs are made in random order, and the results are shown in Table P14.1. Analyze this experiment, assuming all three factors are fixed.

Table P14.1

	Machine 1			Machine 2			Machine 3		
Station	1	2	3	1	2	3	1	2	3
Power	34.1	33.7	36.2	31.1	33.1	32.8	32.9	33.8	33.6
Setting	30.3	34.9	36.8	33.5	34.7	35.1	33.0	33.4	32.8
1	31.6	35.0	37.1	34.0	33.9	34.3	33.1	32.8	31.7
Power	24.3	28.1	25.7	24.1	24.1	26.0	24.2	23.2	24.7
Setting	26.3	29.3	26.1	25.0	25.1	27.1	26.1	27.4	22.0
2	27.1	28.6	24.9	26.3	27.9	23.9	25.3	28.0	24.8

The linear model is $y_{ijkl} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_{k(j)} + (\tau\gamma)_{ik(j)} + \varepsilon_{(ijk)l}$

Minitab Output

```
ANOVA: Yield versus Machine, Power, Station

Factor             Type Levels Values
Machine           fixed      3     1     2     3
Power             fixed      2     1     2
Station(Machine)  fixed      3     1     2     3

Analysis of Variance for Yield

Source                    DF         SS         MS       F      P
Machine                    2     21.436     10.718    6.25  0.005
Power                      1    845.698    845.698  492.96  0.000
Station(Machine)           6     33.583      5.597    3.26  0.012
Machine*Power              2      0.383      0.191    0.11  0.895
Power*Station(Machine)     6     29.208      4.868    2.84  0.023
Error                     36     61.760      1.716
Total                     53    992.068

Source                    Variance Error Expected Mean Square for Each Term
                         component term (using restricted model)
 1 Machine                           6   (6) + 18Q[1]
 2 Power                             6   (6) + 27Q[2]
 3 Station(Machine)                  6   (6) + 6Q[3]
 4 Machine*Power                     6   (6) + 9Q[4]
 5 Power*Station(Machine)            6   (6) + 3Q[5]
 6 Error                     1.716       (6)
```

14.14. Suppose that in Problem 14.13 a large number of power settings could have been used and that the two selected for the experiment were chosen randomly. Obtain the expected mean squares for this situation assuming the restricted form of the mixed model and modify the previous analysis appropriately.

The analysis of variance and the expected mean squares can be obtained from Minitab as follows:

Minitab Output

```
ANOVA: Yield versus Machine, Power, Station

Factor             Type Levels Values
Machine           fixed      3     1     2     3
Power            random      2     1     2
Station(Machine)  fixed      3     1     2     3

Analysis of Variance for Yield

Source                    DF         SS         MS       F      P
Machine                    2     21.436     10.718   56.03  0.018
Power                      1    845.698    845.698  492.96  0.000
Station(Machine)           6     33.583      5.597    1.15  0.435
Machine*Power              2      0.383      0.191    0.11  0.895
Power*Station(Machine)     6     29.208      4.868    2.84  0.023
Error                     36     61.760      1.716
Total                     53    992.068

Source                    Variance Error Expected Mean Square for Each Term
                         component term (using restricted model)
 1 Machine                           4   (6) + 9(4) + 18Q[1]
 2 Power                   31.2586   6   (6) + 27(2)
 3 Station(Machine)                  5   (6) + 3(5) + 6Q[3]
 4 Machine*Power           -0.1694   6   (6) + 9(4)
 5 Power*Station(Machine)   1.0508   6   (6) + 3(5)
 6 Error                    1.7156       (6)
```

14.15. Reanalyze the experiment in Problem 14.14 assuming the unrestricted form of the mixed model. You may use a computer software program to do this. Comment on any differences between the restricted and unrestricted model analysis and conclusions.

Minitab Output

```
ANOVA: Yield versus Machine, Power, Station

Factor            Type Levels Values
Machine          fixed      3     1     2     3
Power           random      2     1     2
Station(Machine)  fixed      3     1     2     3

Analysis of Variance for Yield

Source                    DF         SS         MS       F      P
Machine                    2     21.436     10.718   56.03  0.018
Power                      1    845.698    845.698 4420.88  0.000
Station(Machine)           6     33.583      5.597    1.15  0.435
Machine*Power              2      0.383      0.191    0.04  0.962
Power*Station(Machine)     6     29.208      4.868    2.84  0.023
Error                     36     61.760      1.716
Total                     53    992.068

Source                    Variance Error Expected Mean Square for Each Term
                         component term (using unrestricted model)
 1 Machine                           4   (6) + 3(5) + 9(4) + Q[1,3]
 2 Power                   31.3151   4   (6) + 3(5) + 9(4) + 27(2)
 3 Station(Machine)                  5   (6) + 3(5) + Q[3]
 4 Machine*Power           -0.5196   5   (6) + 3(5) + 9(4)
 5 Power*Station(Machine)   1.0508   6   (6) + 3(5)
 6 Error                    1.7156       (6)
```

There are differences between several of the expected mean squares. However, the conclusions that could be drawn do not differ in any meaningful way from the restricted model analysis.

14.16. A structural engineer is studying the strength of aluminum alloy purchased from three vendors. Each vendor submits the alloy in standard-sized bars of 1.0, 1.5, or 2.0 inches. The processing of different sizes of bar stock from a common ingot involves different forging techniques, and so this factor may be important. Furthermore, the bar stock if forged from ingots made in different heats. Each vendor submits two tests specimens of each size bar stock from the three heats. The resulting strength data is shown in Table P14.2. Analyze the data, assuming that vendors and bar size are fixed and heats are random. Use the restricted form of the mixed model.

Table P14.2

		Vendor 1			Vendor 2			Vendor 3		
	Heat	1	2	3	1	2	3	1	2	3
Bar	1 inch	1.230	1.346	1.235	1.301	1.346	1.315	1.247	1.275	1.324
Size		1.259	1.400	1.206	1.263	1.392	1.320	1.296	1.268	1.315
	1 ½	1.316	1.329	1.250	1.274	1.384	1.346	1.273	1.260	1.392
	inch	1.300	1.362	1.239	1.268	1.375	1.357	1.264	1.265	1.364
	2 inch	1.287	1.346	1.273	1.247	1.362	1.336	1.301	1.280	1.319
		1.292	1.382	1.215	1.215	1.328	1.342	1.262	1.271	1.323

$$y_{ijkl} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \gamma_{k(j)} + (\tau\gamma)_{ik(j)} + \varepsilon_{(ijk)l}$$

Minitab Output

```
ANOVA: Strength versus Vendor, Bar Size, Heat

Factor         Type Levels Values
Vendor        fixed      3     1     2     3
```

```
Heat(Vendor) random      3     1     2     3
Bar Size      fixed      3   1.0   1.5   2.0

Analysis of Variance for Strength

Source                  DF         SS         MS       F      P
Vendor                   2  0.0088486  0.0044243    0.26  0.776
Heat(Vendor)             6  0.1002093  0.0167016   41.32  0.000
Bar Size                 2  0.0025263  0.0012631    1.37  0.290
Vendor*Bar Size          4  0.0023754  0.0005939    0.65  0.640
Bar Size*Heat(Vendor)   12  0.0110303  0.0009192    2.27  0.037
Error                   27  0.0109135  0.0004042
Total                   53  0.1359034

Source                   Variance Error Expected Mean Square for Each Term
                        component term (using restricted model)
 1 Vendor                           2   (6) + 6(2) + 18Q[1]
 2 Heat(Vendor)            0.00272  6   (6) + 6(2)
 3 Bar Size                         5   (6) + 2(5) + 18Q[3]
 4 Vendor*Bar Size                  5   (6) + 2(5) + 6Q[4]
 5 Bar Size*Heat(Vendor)   0.00026  6   (6) + 2(5)
 6 Error                   0.00040      (6)
```

14.17. Rework Problem 14.16 using the unrestricted form of the mixed model. You can use a computer software program to do this. Comment on any differences between the restricted and unrestricted model analysis and conclusions.

Minitab Output

ANOVA: Strength versus Vendor, Bar Size, Heat

```
Factor         Type Levels Values
Vendor        fixed      3     1     2     3
Heat(Vendor) random      3     1     2     3
Bar Size      fixed      3   1.0   1.5   2.0

Analysis of Variance for Strength

Source                  DF         SS         MS       F      P
Vendor                   2  0.0088486  0.0044243    0.26  0.776
Heat(Vendor)             6  0.1002093  0.0167016   18.17  0.000
Bar Size                 2  0.0025263  0.0012631    1.37  0.290
Vendor*Bar Size          4  0.0023754  0.0005939    0.65  0.640
Bar Size*Heat(Vendor)   12  0.0110303  0.0009192    2.27  0.037
Error                   27  0.0109135  0.0004042
Total                   53  0.1359034

Source                   Variance Error Expected Mean Square for Each Term
                        component term (using unrestricted model)
 1 Vendor                           2   (6) + 2(5) + 6(2) + Q[1,4]
 2 Heat(Vendor)            0.00263  5   (6) + 2(5) + 6(2)
 3 Bar Size                         5   (6) + 2(5) + Q[3,4]
 4 Vendor*Bar Size                  5   (6) + 2(5) + Q[4]
 5 Bar Size*Heat(Vendor)   0.00026  6   (6) + 2(5)
 6 Error                   0.00040      (6)
```

There are some differences in the expected mean squares. However, the conclusions do not differ from those of the restricted model analysis.

14.18. Suppose that in Problem 14.16 the bar stock may be purchased in many sizes and that the three sizes actually used in the experiment were selected randomly. Obtain the expected mean squares for this situation and modify the previous analysis appropriately. Use the restricted form of the mixed model.

Minitab Output

```
ANOVA: Strength versus Vendor, Bar Size, Heat

Factor         Type Levels Values
Vendor        fixed      3     1     2     3
Heat(Vendor) random      3     1     2     3
Bar Size     random      3   1.0   1.5   2.0

Analysis of Variance for Strength

Source                   DF         SS         MS       F      P
Vendor                    2  0.0088486  0.0044243    0.27  0.772 x
Heat(Vendor)              6  0.1002093  0.0167016   18.17  0.000
Bar Size                  2  0.0025263  0.0012631    1.37  0.290
Vendor*Bar Size           4  0.0023754  0.0005939    0.65  0.640
Bar Size*Heat(Vendor)    12  0.0110303  0.0009192    2.27  0.037
Error                    27  0.0109135  0.0004042
Total                    53  0.1359034

x Not an exact F-test.

Source                   Variance Error Expected Mean Square for Each Term
                        component term (using restricted model)
 1 Vendor                           *   (6) + 2(5) + 6(4) + 6(2) + 18Q[1]
 2 Heat(Vendor)            0.00263   5   (6) + 2(5) + 6(2)
 3 Bar Size                0.00002   5   (6) + 2(5) + 18(3)
 4 Vendor*Bar Size        -0.00005   5   (6) + 2(5) + 6(4)
 5 Bar Size*Heat(Vendor)  0.00026   6   (6) + 2(5)
 6 Error                   0.00040       (6)
* Synthesized Test.

Error Terms for Synthesized Tests

Source                    Error DF  Error MS  Synthesis of Error MS
 1 Vendor                     5.75 0.0163762   (2) + (4) - (5)
```

Notice that a Satterthwaite type test is used for vendor.

14.19. Steel in normalized by heating above the critical temperature, soaking, and then air cooling. This process increases the strength of the steel, refines the grain, and homogenizes the structure. An experiment is performed to determine the effect of temperature and heat treatment time on the strength of normalized steel. Two temperatures and three times are selected. The experiment is performed by heating the oven to a randomly selected temperature and inserting three specimens. After 10 minutes one specimen is removed, after 20 minutes the second specimen is removed, and after 30 minutes the final specimen is removed. Then the temperature is changed to the other level and the process is repeated. Four shifts are required to collect the data, which are shown below. Analyze the data and draw conclusions, assume both factors are fixed.

		Temperature (F)	
Shift	Time(minutes)	1500	1600
1	10	63	89
	20	54	91
	30	61	62
2	10	50	80
	20	52	72
	30	59	69
3	10	48	73
	20	74	81
	30	71	69
4	10	54	88
	20	48	92
	30	59	64

This is a split-plot design. Shifts correspond to blocks, temperature is the whole plot treatment, and time is the subtreatments (in the subplot or split-plot part of the design). The expected mean squares and analysis of variance are shown below. The following Minitab Output has been modified to display the results of the split-plot analysis. Minitab will calculate the sums of squares correctly, but the expected mean squares and the statistical tests are not, in general, correct. Notice that the Error term in the analysis of variance is actually the three factor interaction.

Minitab Output

ANOVA: Strength versus Shift, Temperature, Time

```
Factor     Type Levels Values
Shift    random      4     1     2     3     4
Temperat  fixed      2  1500  1600
Time      fixed      3    10    20    30

Analysis of Variance for Strength
                                                Standard     Split Plot
Source            DF          SS          MS       F      P       F      P
Shift              3      145.46       48.49    1.19  0.390
Temperat           1     2340.38     2340.38   29.20  0.012   29.21  0.012
Shift*Temperat     3      240.46       80.15    1.97  0.220
Time               2      159.25       79.63    1.00  0.422    1.00  0.422
Shift*Time         6      478.42       79.74    1.96  0.217
Temperat*Time      2      795.25      397.63    9.76  0.013    9.76  0.013
Error              6      244.42       40.74
Total             23     4403.63

Source            Variance Error Expected Mean Square for Each Term
                 component term (using restricted model)
 1 Shift             1.292   7   (7) + 6(1)
 2 Temperat                  3   (7) + 3(3) + 12Q[2]
 3 Shift*Temperat   13.139   7   (7) + 3(3)
 4 Time                      5   (7) + 2(5) + 8Q[4]
 5 Shift*Time       19.500   7   (7) + 2(5)
 6 Temperat*Time             7   (7) + 4Q[6]
 7 Error            40.736       (7)
```

14.20. An experiment is designed to study pigment dispersion in paint. Four different mixes of a particular pigment are studied. The procedure consists of preparing a particular mix and then applying that mix to a panel by three application methods (brushing, spraying, and rolling). The response measured is the percentage reflectance of the pigment. Three days are required to run the experiment, and the data obtained follow. Analyze the data and draw conclusions, assuming that mixes and application methods are fixed.

		Mix			
Day	App Method	1	2	3	4
1	1	64.5	66.3	74.1	66.5
	2	68.3	69.5	73.8	70.0
	3	70.3	73.1	78.0	72.3
2	1	65.2	65.0	73.8	64.8
	2	69.2	70.3	74.5	68.3
	3	71.2	72.8	79.1	71.5
3	1	66.2	66.5	72.3	67.7
	2	69.0	69.0	75.4	68.6
	3	70.8	74.2	80.1	72.4

This is a split plot design. Days correspond to blocks, mix is the whole plot treatment, and method is the subtreatment (in the subplot or split plot part of the design). The following Minitab Output has been modified to display the results of the split-plot analysis. Minitab will calculate the sums of squares correctly, but the expected mean squares and the statistical tests are not, in general, correct. Notice that the Error term in the analysis of variance is actually the three factor interaction.

Minitab Output

ANOVA: Reflectance versus Day, Mix, Method

```
Factor     Type Levels Values
Day      random      3     1     2     3
Mix       fixed      4     1     2     3     4
Method    fixed      3     1     2     3

Analysis of Variance for Reflecta
                                            Standard     Split Plot
Source        DF          SS          MS       F      P       F      P
Day            2       2.042       1.021    1.39  0.285
Mix            3     307.479     102.493  135.77  0. 000 135.75  0.000
Day*Mix        6       4.529       0.755    1.03  0.451
Method         2     222.095     111.047  226.24  0.000  226.16  0.000
Day*Method     4       1.963       0.491    0.67  0.625
Mix*Method     6      10.036       1.673    2.28  0.105    2.28  0.105
Error         12       8.786       0.732
Total         35     556.930

Source        Variance Error Expected Mean Square for Each Term
             component term (using restricted model)
 1 Day         0.02406   7   (7) + 12(1)
 2 Mix                   3   (7) + 3(3) + 9Q[2]
 3 Day*Mix     0.00759   7   (7) + 3(3)
 4 Method                5   (7) + 4(5) + 12Q[4]
 5 Day*Method -0.06032   7   (7) + 4(5)
 6 Mix*Method            7   (7) + 3Q[6]
 7 Error       0.73213       (7)
```

14.21. Repeat Problem 14.20, assuming that the mixes are random and the application methods are fixed.

The F-tests are the same as those in Problem 13-20. The following Minitab Output has been edited to display the results of the split-plot analysis. Minitab will calculate the sums of squares correctly, but the expected mean squares and the statistical tests are not, in general, correct. Again, the Error term in the analysis of variance is actually the three factor interaction.

Minitab Output

ANOVA: Reflectance versus Day, Mix, Method

```
Factor     Type Levels Values
Day      random      3     1     2     3
Mix      random      4     1     2     3     4
Method    fixed      3     1     2     3

Analysis of Variance for Reflecta
                                            Standard        Split Plot
Source        DF          SS          MS       F      P          F      P
Day            2       2.042       1.021    1.35  0.328
Mix            3     307.479     102.493  135.77  0.000     135.75  0.000
Day*Mix        6       4.529       0.755    1.03  0.451
Method         2     222.095     111.047   77.58  0.001 x   226.16  0.000
Day*Method     4       1.963       0.491    0.67  0.625
Mix*Method     6      10.036       1.673    2.28  0.105       2.28  0.105
Error         12       8.786       0.732
Total         35     556.930

x Not an exact F-test.

Source        Variance Error Expected Mean Square for Each Term
             component term (using restricted model)
 1 Day          0.0222   3   (7) + 3(3) + 12(1)
```

```
 2 Mix         11.3042   3   (7) + 3(3) + 9(2)
 3 Day*Mix      0.0076   7   (7) + 3(3)
 4 Method                *   (7) + 3(6) + 4(5) + 12Q[4]
 5 Day*Method  -0.0603   7   (7) + 4(5)
 6 Mix*Method   0.3135   7   (7) + 3(6)
 7 Error        0.7321       (7)

* Synthesized Test.

Error Terms for Synthesized Tests

Source         Error DF  Error MS  Synthesis of Error MS
 4 Method          3.59     1.431   (5) + (6) - (7)
```

14.22. Consider the split-split-plot design described in Example 14.4. Suppose that this experiment is conducted as described and that the data shown in Table P14.3 are obtained. Analyze the data and draw conclusions.

Table P14.3

		Technician								
		1			2			3		
Blocks	Dose Strength	1	2	3	1	2	3	1	2	3
	Wall Thickness									
1	1	95	71	108	96	70	108	95	70	100
	2	104	82	115	99	84	100	102	81	106
	3	101	85	117	95	83	105	105	84	113
	4	108	85	116	97	85	109	107	87	115
2	1	95	78	110	100	72	104	92	69	101
	2	106	84	109	101	79	102	100	76	104
	3	103	86	116	99	80	108	101	80	109
	4	109	84	110	112	86	109	108	86	113
3	1	96	70	107	94	66	100	90	73	98
	2	105	81	106	100	84	101	97	75	100
	3	106	88	112	104	87	109	100	82	104
	4	113	90	117	121	90	117	110	91	112
4	1	90	68	109	98	68	106	98	72	101
	2	100	84	112	102	81	103	102	78	105
	3	102	85	115	100	85	110	105	80	110
	4	114	88	118	118	85	116	110	95	120

Using the computer output, the F-ratios were calculated by hand using the expected mean squares found in Table 14.18. The following Minitab Output has been edited to display the results of the split-plot analysis. Minitab will calculate the sums of squares correctly, but the expected mean squares and the statistical tests are not, in general, correct. Notice that the Error term in the analysis of variance is actually the four factor interaction.

Minitab Output

ANOVA: Time versus Day, Tech, Dose, Thick

```
Factor     Type Levels Values
Day      random      4     1     2     3     4
Tech      fixed      3     1     2     3
Dose      fixed      3     1     2     3
Thick     fixed      4     1     2     3     4

Analysis of Variance for Time
                                                Standard     Split Plot
Source             DF         SS          MS       F      P       F      P
```

Day	3	48.41	16.14	3.38	0.029		
Tech	2	248.35	124.17	4.62	0.061	4.62	0.061
Day*Tech	6	161.15	26.86	5.62	0.000		
Dose	2	20570.06	10285.03	550.44	0.000	550.30	0.000
Day*Dose	6	112.11	18.69	3.91	0.004		
Tech*Dose	4	125.94	31.49	3.32	0.048	3.32	0.048
Day*Tech*Dose	12	113.89	9.49	1.99	0.056		
Thick	3	3806.91	1268.97	36.47	0.000	36.48	0.000
Day*Thick	9	313.12	34.79	7.28	0.000		
Tech*Thick	6	126.49	21.08	2.26	0.084	2.26	0.084
Day*Tech*Thick	18	167.57	9.31	1.95	0.044		
Dose*Thick	6	402.28	67.05	17.13	0.000	17.15	0.000
Day*Dose*Thick	18	70.44	3.91	0.82	0.668		
Tech*Dose*Thick	12	205.89	17.16	3.59	0.001	3.59	0.001
Error	36	172.06	4.78				
Total	143	26644.66					

Source	Variance component	Error term	Expected Mean Square for Each Term (using restricted model)
1 Day	0.3155	15	(15) + 36(1)
2 Tech		3	(15) + 12(3) + 48Q[2]
3 Day*Tech	1.8400	15	(15) + 12(3)
4 Dose		5	(15) + 12(5) + 48Q[4]
5 Day*Dose	1.1588	15	(15) + 12(5)
6 Tech*Dose		7	(15) + 4(7) + 16Q[6]
7 Day*Tech*Dose	1.1779	15	(15) + 4(7)
8 Thick		9	(15) + 9(9) + 36Q[8]
9 Day*Thick	3.3346	15	(15) + 9(9)
10 Tech*Thick		11	(15) + 3(11) + 12Q[10]
11 Day*Tech*Thick	1.5100	15	(15) + 3(11)
12 Dose*Thick		13	(15) + 3(13) + 12Q[12]
13 Day*Dose*Thick	-0.2886	15	(15) + 3(13)
14 Tech*Dose*Thick		15	(15) + 4Q[14]
15 Error	4.7793		(15)

14.23. Rework Problem 14.22, assuming that the dosage strengths are chosen at random. Use the restricted form of the mixed model.

The following Minitab Output has been edited to display the results of the split-plot analysis. Minitab will calculate the sums of squares correctly, but the expected mean squares and the statistical tests are not, in general, correct. Again, the Error term in the analysis of variance is actually the four factor interaction.

Minitab Output

ANOVA: Time versus Day, Tech, Dose, Thick

Factor	Type	Levels	Values			
Day	random	4	1	2	3	4
Tech	fixed	3	1	2	3	
Dose	random	3	1	2	3	
Thick	fixed	4	1	2	3	4

Analysis of Variance for Time

Source	DF	SS	MS	Standard F	Standard P		Split Plot F	Split Plot P
Day	3	48.41	16.14	0.86	0.509			
Tech	2	248.35	124.17	2.54	0.155		4.62	0.061
Day*Tech	6	161.15	26.86	2.83	0.059			
Dose	2	20570.06	10285.03	550.44	0.000		550.30	0.000
Day*Dose	6	112.11	18.69	3.91	0.004			
Tech*Dose	4	125.94	31.49	3.32	0.048		3.32	0.048
Day*Tech*Dose	12	113.89	9.49	1.99	0.056			
Thick	3	3806.91	1268.97	12.96	0.001	x	36.48	0.000
Day*Thick	9	313.12	34.79	8.89	0.000			
Tech*Thick	6	126.49	21.08	0.97	0.475	x	2.26	0.084
Day*Tech*Thick	18	167.57	9.31	1.95	0.044			
Dose*Thick	6	402.28	67.05	17.13	0.000		17.15	0.000
Day*Dose*Thick	18	70.44	3.91	0.82	0.668			

```
Tech*Dose*Thick    12     205.89      17.16    3.59  0.001       3.59  0.001
Error              36     172.06       4.78
Total             143   26644.66

x Not an exact F-test.

Source              Variance Error Expected Mean Square for Each Term
                   component term (using restricted model)
 1 Day               -0.071   5   (15) + 12(5) + 36(1)
 2 Tech                       *   (15) + 4(7) + 16(6) + 12(3) + 48Q[2]
 3 Day*Tech           1.447   7   (15) + 4(7) + 12(3)
 4 Dose             213.882   5   (15) + 12(5) + 48(4)
 5 Day*Dose           1.159  15   (15) + 12(5)
 6 Tech*Dose          1.375   7   (15) + 4(7) + 16(6)
 7 Day*Tech*Dose      1.178  15   (15) + 4(7)
 8 Thick                      *   (15) + 3(13) + 12(12) + 9(9) + 36Q[8]
 9 Day*Thick          3.431  13   (15) + 3(13) + 9(9)
10 Tech*Thick                 *   (15) + 4(14) + 3(11) + 12Q[10]
11 Day*Tech*Thick     1.510  15   (15) + 3(11)
12 Dose*Thick         5.261  13   (15) + 3(13) + 12(12)
13 Day*Dose*Thick    -0.289  15   (15) + 3(13)
14 Tech*Dose*Thick    3.095  15   (15) + 4(14)
15 Error              4.779       (15)

* Synthesized Test.

Error Terms for Synthesized Tests

Source              Error DF  Error MS  Synthesis of Error MS
 2 Tech                 6.35     48.85   (3) + (6) - (7)
 8 Thick               10.84     97.92   (9) + (12) - (13)
10 Tech*Thick          15.69     21.69   (11) + (14) - (15)
```

There are no exact tests on technicians β_j, dosage strengths γ_k, wall thickness δ_h, or the technician x wall thickness interaction $(\beta\delta)_{jh}$. The approximate F-tests are as follows:

H_0: $\beta_j = 0$

$$F = \frac{MS_B + MS_{ABC}}{MS_{AB} + MS_{BC}} = \frac{124.174 + 9.491}{26.859 + 31.486} = 2.291$$

$$p = \frac{(MS_B + MS_{ABC})^2}{\frac{MS_B^2}{2} + \frac{MS_{ABC}^2}{12}} = \frac{(124.174 + 9.491)^2}{\frac{124.174^2}{2} + \frac{9.491^2}{12}} = 2.315$$

$$q = \frac{(MS_{AB} + MS_{BC})^2}{\frac{MS_{AB}^2}{6} + \frac{MS_{BC}^2}{4}} = \frac{(26.859 + 31.486)^2}{\frac{26.859^2}{6} + \frac{31.486^2}{4}} = 9.248$$

Do not reject H_0: $\beta_j = 0$

H_0: $\gamma_k = 0$

$$F = \frac{MS_C + MS_{ACD}}{MS_{CD} + MS_{AD}} = \frac{10285.028 + 3.914}{67.046 + 34.791} = 101.039$$

$$p = \frac{(MS_C + MS_{ACD})^2}{\frac{MS_C^2}{2} + \frac{MS_{ACD}^2}{18}} = \frac{(10285.028 + 3.914)^2}{\frac{10285.028^2}{2} + \frac{3.914^2}{18}} = 2.002$$

$$q = \frac{(MS_{CD} + MS_{AD})^2}{\frac{MS_{CD}^2}{6} + \frac{MS_{AD}^2}{9}} = \frac{(67.046 + 34.791)^2}{\frac{67.046^2}{6} + \frac{34.791^2}{9}} = 11.736$$

Reject H_0: $\gamma_k = 0$

H_0: $\delta_h = 0$

$$F = \frac{MS_D + MS_{ACD}}{MS_{CD} + MS_{AD}} = \frac{1268.970 + 3.914}{67.046 + 34.791} = 12.499$$

$$p = \frac{(MS_D + MS_{ACD})^2}{\frac{MS_D^2}{3} + \frac{MS_{ACD}^2}{18}} = \frac{(1268.970 + 3.914)^2}{\frac{1268.970^2}{3} + \frac{3.914^2}{18}} = 3.019$$

$$q = \frac{(MS_{CD} + MS_{AD})^2}{\frac{MS_{CD}^2}{6} + \frac{MS_{AD}^2}{9}} = \frac{(67.046 + 34.791)^2}{\frac{67.046^2}{6} + \frac{34.791^2}{9}} = 11.736$$

Reject H_0: $\delta_h = 0$

H_0: $(\beta\delta)_{jh} = 0$

$$F = \frac{MS_{BD} + MS_{ABCD}}{MS_{BCD} + MS_{ABD}} = \frac{21.081 + 4.779}{17.157 + 9.309} = 0.977$$

$F<1$, Do not reject H_0: $(\beta\delta)_{jh} = 0$

14.24. Suppose that in Problem 14.22 four technicians had been used. Assuming that all the factors are fixed, how many blocks should be run to obtain an adequate number of degrees of freedom on the test for differences among technicians?

The number of degrees of freedom for the test is (a-1)(4-1)=3(a-1), where a is the number of blocks used.

Number of Blocks (a)	DF for test
2	3
3	6
4	9
5	12

At least three blocks should be run, but four would give a better test.

14.25. Consider the experiment described in Example 14.4. Demonstrate how the order in which the treatments combinations are run would be determined if this experiment were run as (a) a split-split-plot, (b) a split-plot, (c) a factorial design in a randomized block, and (d) a completely randomized factorial design.

(a) Randomization for the split-split plot design is described in Example 14.4.

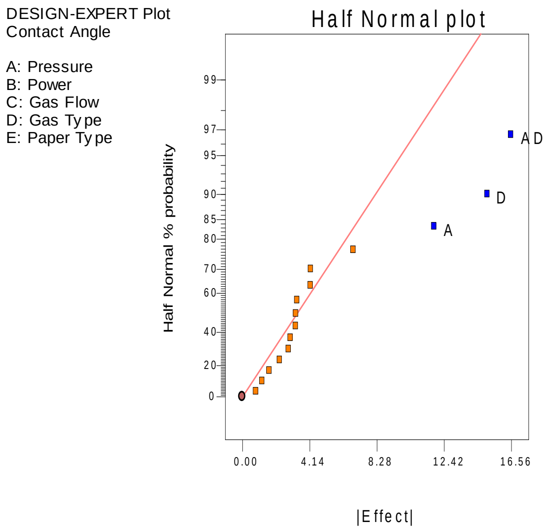
(b) In the split-plot, within a block, the technicians would be the main treatment and within a block-technician plot, the 12 combinations of dosage strength and wall thickness would be run in random order. The design would be a two-factor factorial in a split-plot.
(c) To run the design in a randomized block, the 36 combinations of technician, dosage strength, and wall thickness would be run in random order within each block. The design would be a three factor factorial in a randomized block.
(d) The blocks would be considered as replicates, and all 144 observations would be 4 replicates of a three factor factorial.

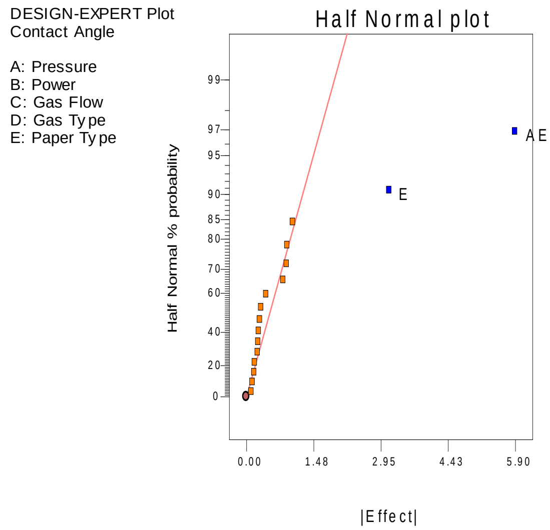
14.26. An article in *Quality Engineering* ("Quality Quandaries: Two-Level Factorials Run as Split-Plot Experiments", Bisgaard, et al, Vol. 8, No. 4, pp. 705-708, 1996) describes a 2^5 factorial experiment on a plasma process focused on making paper more susceptible to ink. Four of the factors (*A*-*D*) are difficult to change from run-to-run, so the experimenters set up the reactor at the eight sets of conditions specified by the low and high levels of these factors, and then processed the two paper types (factor *E*) together. The placement of the paper specimens in the reactors (right versus left) was randomized. This produces a split-plot design with *A*-*D* as the whole-plot factors and factor *E* as the subplot factor. The data from this experiment are shown in Table P14.4. Analyze the data from this experiment and draw conclusions.

Table P14.4

Standard Order	Run Number	*A* = Pressure	*B* = Power	*C* = Gas Flow	*D* = Gas Type	*E* = Paper Type	*y* Contact Angle
1	23	-1	-1	-1	Oxygen	E1	48.6
2	3	+1	-1	-1	Oxygen	E1	41.2
3	11	-1	+1	-1	Oxygen	E1	55.8
4	29	+1	+1	-1	Oxygen	E1	53.5
5	1	-1	-1	+1	Oxygen	E1	37.6
6	15	+1	-1	+1	Oxygen	E1	47.2
7	27	-1	+1	+1	Oxygen	E1	47.2
8	25	+1	+1	+1	Oxygen	E1	48.7
9	19	-1	-1	-1	SiCl4	E1	5
10	5	+1	-1	-1	SiCl4	E1	56.8
11	9	-1	+1	-1	SiCl4	E1	25.6
12	31	+1	+1	-1	SiCl4	E1	41.8
13	13	-1	-1	+1	SiCl4	E1	13.3
14	7	+1	-1	+1	SiCl4	E1	47.5
15	21	-1	+1	+1	SiCl4	E1	11.3
16	17	+1	+1	+1	SiCl4	E1	49.5
17	24	-1	-1	-1	Oxygen	E2	57
18	4	+1	-1	-1	Oxygen	E2	38.2
19	12	-1	+1	-1	Oxygen	E2	62.9
20	30	+1	+1	-1	Oxygen	E2	51.3
21	2	-1	-1	+1	Oxygen	E2	43.5
22	16	+1	-1	+1	Oxygen	E2	44.8
23	28	-1	+1	+1	Oxygen	E2	54.6
24	26	+1	+1	+1	Oxygen	E2	44.4
25	20	-1	-1	-1	SiCl4	E2	18.1
26	6	+1	-1	-1	SiCl4	E2	56.2
27	10	-1	+1	-1	SiCl4	E2	33
28	32	+1	+1	-1	SiCl4	E2	37.8
29	14	-1	-1	+1	SiCl4	E2	23.7
30	8	+1	-1	+1	SiCl4	E2	43.2
31	22	-1	+1	+1	SiCl4	E2	23.9
32	18	+1	+1	+1	SiCl4	E2	48.2

Half normal probability plots of the effects for both the whole plot with factors *A*, *B*, *C*, *D*, and their corresponding interactions, as well as the sub-plot with factor *E* and all interactions involving *E*, are shown below. The analysis of variance is not shown because of the known errors in the calculations; however, the models are also shown below.





Design Expert Output

Response: Contact Angle

Final Equation in Terms of Coded Factors:

Contact Angle =
+40.98
+5.91 * A
-7.55 * D
+1.57 * E
+8.28 * A * D
-2.95 * A * E

Final Equation in Terms of Actual Factors:

Gas Type Oxygen
Paper Type E1
Contact Angle =
+46.96250
+0.58125 * Pressure

Gas Type SiCl4
Paper Type E1
Contact Angle =
+31.86250
+17.14375 * Pressure

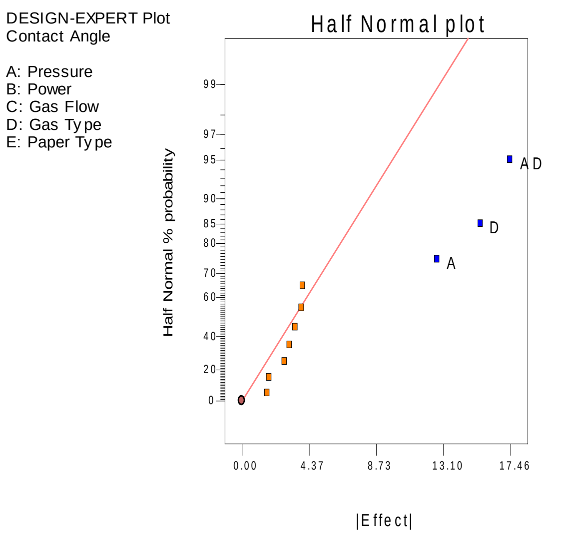
Gas Type Oxygen
Paper Type E2
Contact Angle =
+50.10000
-5.31875 * Pressure

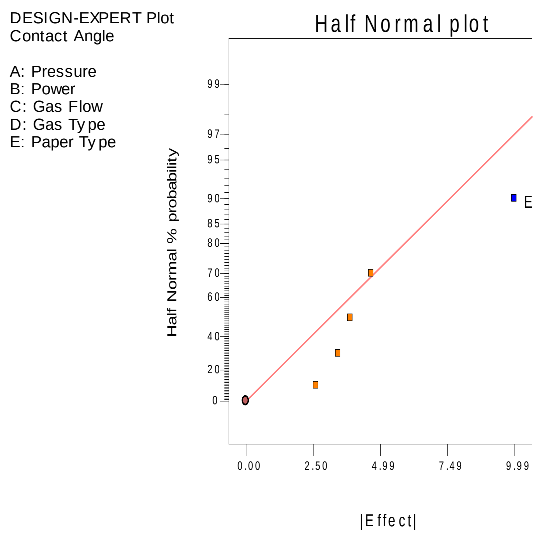
Gas Type SiCl4
Paper Type E2
Contact Angle =
+35.00000
+11.24375 * Pressure

14.27. Reconsider the experiment in problem 14.26. This is a rather large experiment, so suppose that the experimenter had used a 2^{5-1} design instead. Set up the 2^{5-1} design in a split-plot, using the principle fraction. Then select the response data using the information from the full factorial. Analyze the data and draw conclusions. Do they agree with the results of Problem 14.26?

Standard Order	Run Number	*A* = Pressure	*B* = Power	*C* = Gas Flow	*D* = Gas Type	*E* = Paper Type	*y* Contact Angle
1	12	-1	-1	-1	Oxygen	E2	57
2	2	+1	-1	-1	Oxygen	E1	41.2
3	6	-1	+1	-1	Oxygen	E1	55.8
4	15	+1	+1	-1	Oxygen	E2	51.3
5	1	-1	-1	+1	Oxygen	E1	37.6
6	8	+1	-1	+1	Oxygen	E2	44.8
7	14	-1	+1	+1	Oxygen	E2	54.6
8	13	+1	+1	+1	Oxygen	E1	48.7
9	10	-1	-1	-1	SiCl4	E1	5
10	3	+1	-1	-1	SiCl4	E2	56.2
11	5	-1	+1	-1	SiCl4	E2	33
12	16	+1	+1	-1	SiCl4	E1	41.8
13	7	-1	-1	+1	SiCl4	E2	23.7
14	4	+1	-1	+1	SiCl4	E1	47.5
15	11	-1	+1	+1	SiCl4	E1	11.3
16	9	+1	+1	+1	SiCl4	E2	48.2

Similar results are found with the half fraction other than the AE interaction is no longer significant and the effect for factor E is larger. The half normal probability plot of effects for the whole and sub-plots are shown below. The resulting model is also shown.





Design Expert Output

Response: Contact Angle

Final Equation in Terms of Coded Factors:

```
Contact Angle  =
       +41.11
        +6.36  * A
        -7.77  * D
        +4.99  * E
        +8.73  * A * D
```

Final Equation in Terms of Actual Factors:

Gas Type Oxygen
Paper Type E1
Contact Angle =
+43.88125
-2.37500 * Pressure

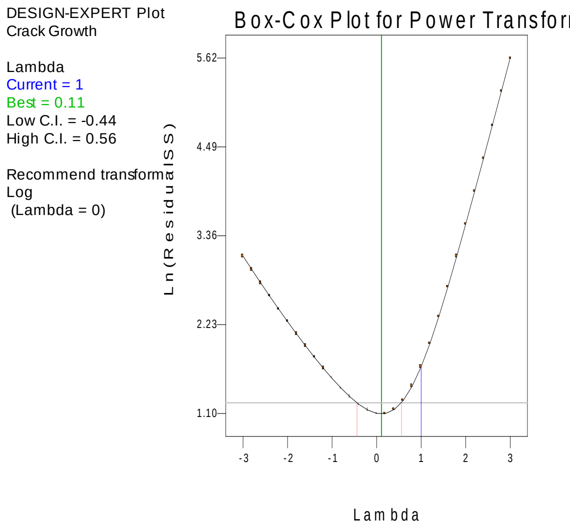
Gas Type SiCl4
Paper Type E1
Contact Angle =
+28.34375
+15.08750 * Pressure

Gas Type Oxygen
Paper Type E2
Contact Angle =
+53.86875
-2.37500 * Pressure

Gas Type SiCl4
Paper Type E2
Contact Angle =
+38.33125
+15.08750 * Pressure

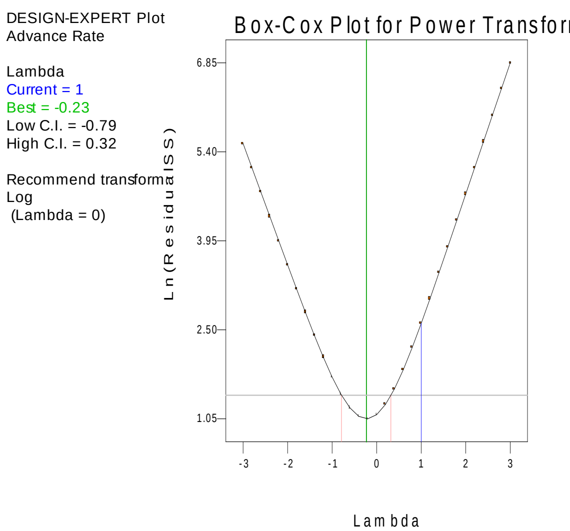
Chapter 15
Other Design and Analysis Topics
Solutions

15.1. Reconsider the experiment in Problem 5.24. Use the Box-Cox procedure to determine if a transformation on the response is appropriate (or useful) in the analysis of the data from this experiment.



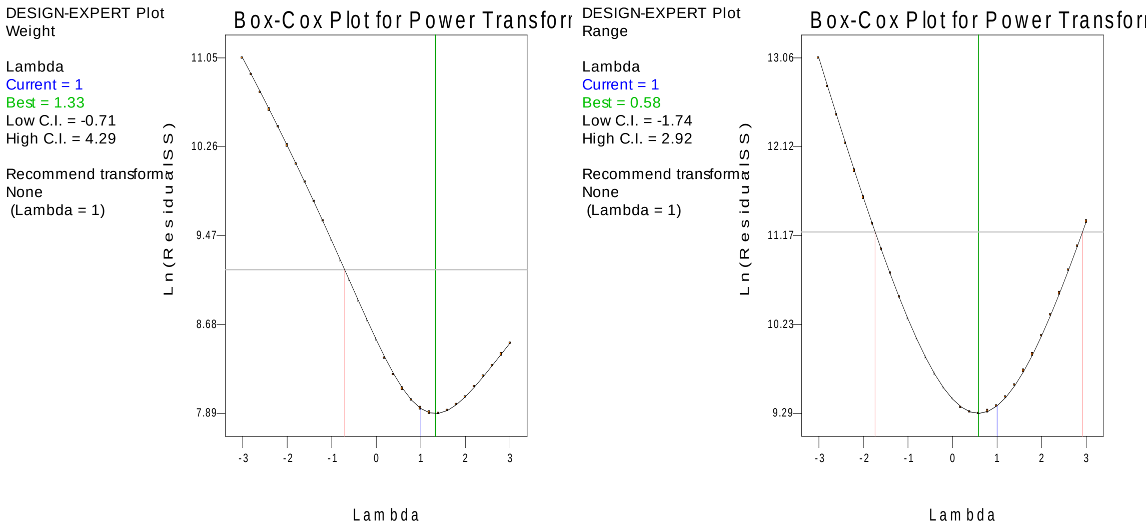
With the value of lambda near zero, and since the confidence interval does not include one, a natural log transformation would be appropriate.

15.2. In Example 6.3 we selected a log transformation for the drill advance rate response. Use the Box-Cox procedure to demonstrate that this is an appropriate data transformation.



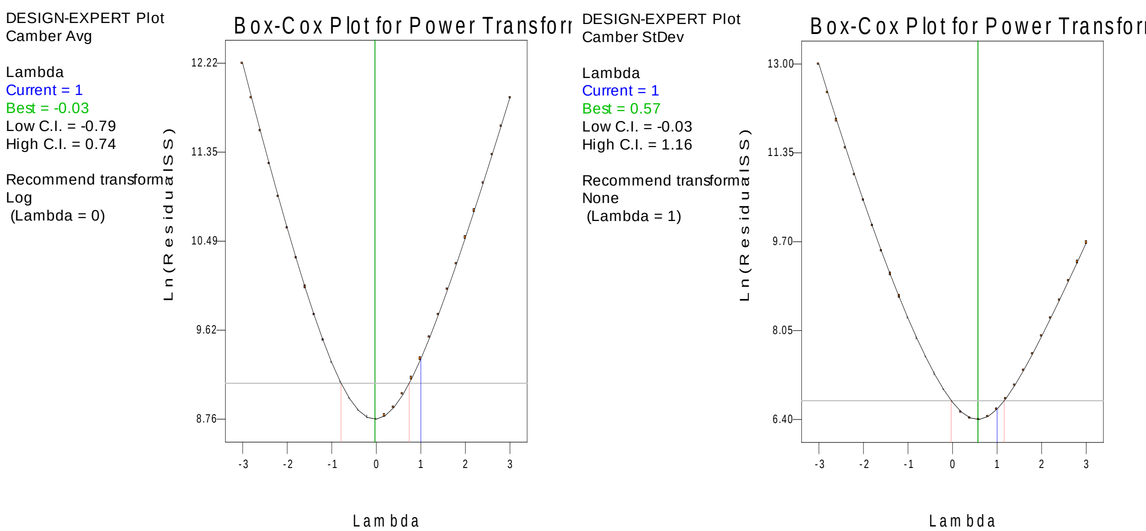
Because the value of lambda is very close to zero, and the confidence interval does not include one, the natural log was the correct transformation chosen for this analysis.

15.3. Reconsider the smelting process experiment in Problem 8.24, where a 2^{6-3} fractional factorial design was used to study the weight of packing material stuck to carbon anodes after baking. Each of the eight runs in the design was replicated three times and both the average weight and the range of the weights at each test combination were treated as response variables. Is there any indication that a transformation is required for either response?



There is no indication that a transformation is required for either response.

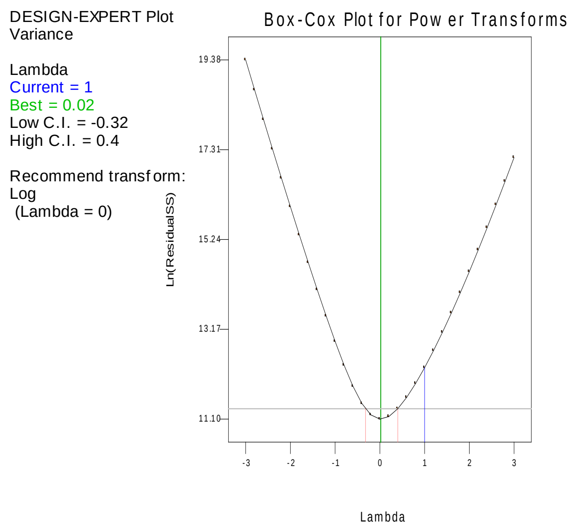
15.4. In Problem 8.25 a replicated fractional factorial design was used to study substrate camber in semiconductor manufacturing. Both the mean and standard deviation of the camber measurements were used as response variables. Is there any indication that a transformation is required for either response?



The Box-Cox plot for the Camber Average suggests a natural log transformation should be applied. This decision is based on the confidence interval for lambda not including one and the point estimate of lambda being very close to zero. With a lambda of approximately 0.5, a square root transformation could be considered for the Camber Standard Deviation; however, the confidence interval indicates that no transformation is needed.

15.5. Reconsider the photoresist experiment in Problem 8.26. Use the variance of the resist thickness at each test combination as the response variable. Is there any indication that a transformation is required?

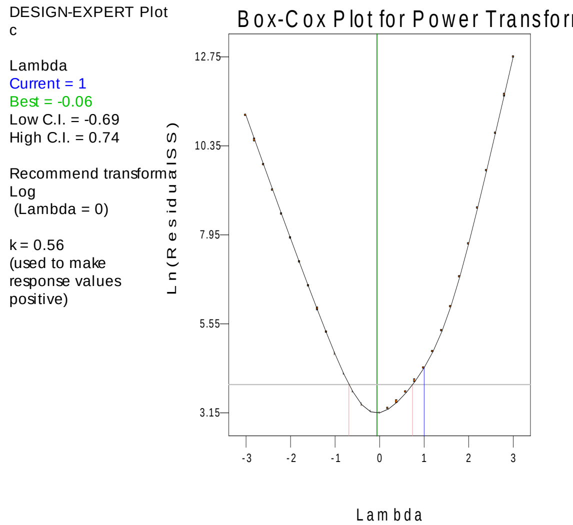
Table P8.4 from Problem 8.26 presents the range, and not the variance. The variance must first be calculated, and the most appropriate model fit to the data. A Box-Cox plot is shown below.



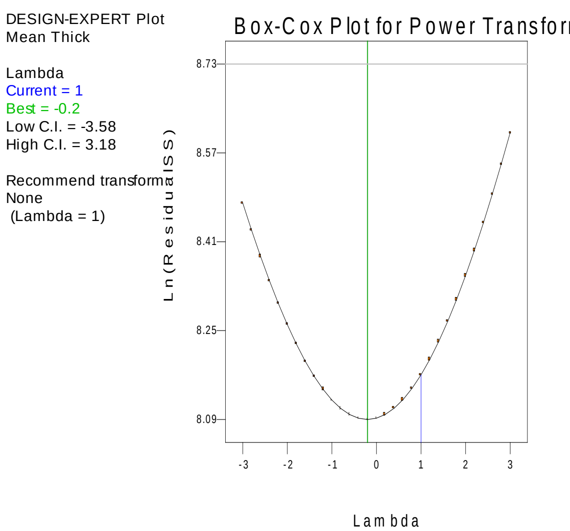
With the point estimate of lambda near zero, and the confidence interval for lambda not inclusive of one, a log transformation would be appropriate. The parameters included in the model affect this plot.

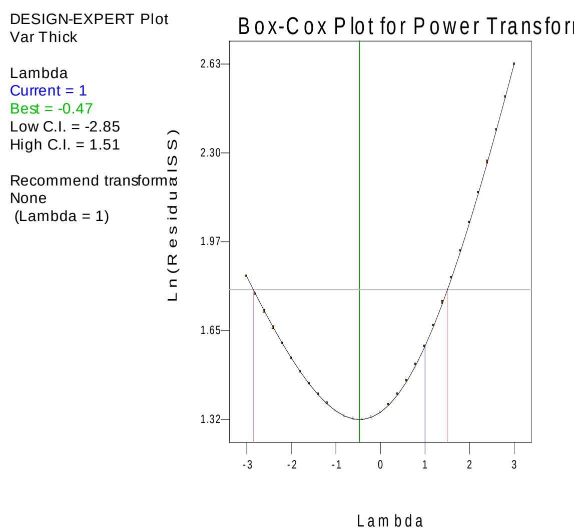
15.6. In the grill defects experiment described in Problem 8.30 a variation of the square root transformation was employed in the analysis of the data. Use the Box-Cox method to determine if this is the appropriate transformation.

The Box-Cox plot is shown below. Because the confidence interval for the minimum lambda does not include one, the decision to use a transformation is correct. Because the lambda point estimate is close to zero, the natural log transformation would be appropriate. This is a stronger transformation than the square root.



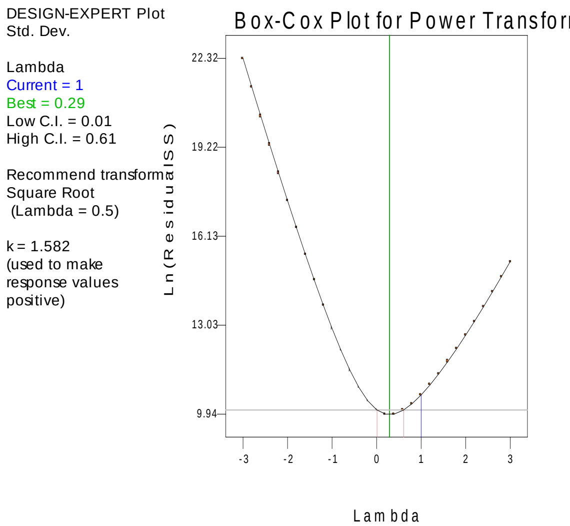
15.7. In the central composite design of Problem 11.14, two responses were obtained, the mean and variance of an oxide thickness. Use the Box-Cox method to investigate the potential usefulness of transformation for both of these responses. Is the log transformation suggested in part (c) of that problem appropriate?





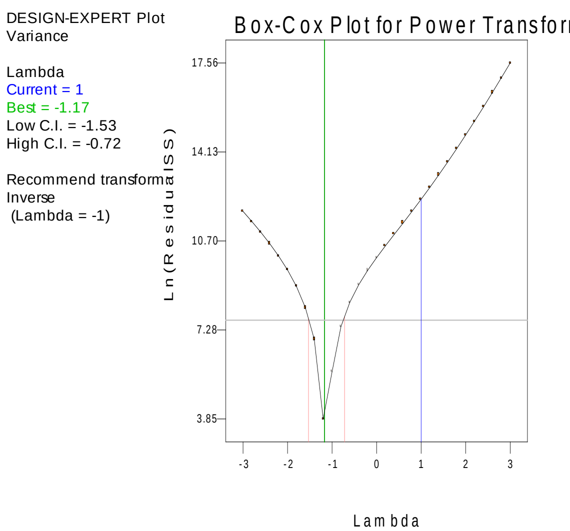
The Box-Cox plot for the Mean Thickness model suggests that a natural log transformation could be applied; however, the confidence interval for lambda includes one. Therefore, a transformation would have a minimal effect. The natural log transformation applied to the Variance of Thickness model appears to be acceptable; however, again the confidence interval for lambda includes one.

15.8. In the 3^3 factorial design of Problem 12.11 one of the responses is a standard deviation. Use the Box-Cox method to investigate the usefulness of transformations for this response. Would your answer change if we used the variance of the response?



Because the confidence interval for lambda does not include one, a transformation should be applied. The square root transformation appears to be acceptable. However, notice that the value of zero is very close to the lower confidence limit, and the minimizing value of lambda is between 0 and 0.5. It is likely that either the natural log or the square root transformation would work reasonably well.

15.9. Problem 12.9 suggests using the $ln(s^2)$ as the response (refer to part b). Does the Box-Cox method indicate that a transformation is appropriate?



Because the confidence interval for lambda does not include one, a transformation should be applied. The confidence interval does not include zero; therefore, the natural log transformation is inappropriate. With the point estimate of lambda at –1.17, the reciprocal transformation is appropriate.

15.10. Myers, Montgomery and Vining (2002) describe an experiment to study spermatozoa survival. The design factors are the amount of sodium citrate, the amount of glycerol, and equilibrium time, each at two levels. The response variable is the number of spermatozoa that survive out of fifty that were tested at each set of conditions. The data are in the following table. Analyze the data from this experiment with logistical regression.

Sodium Citrate	Glycerol	Equilibrium Time	Number Survived
-	-	-	34
+	-	-	20
-	+	-	8
+	+	-	21
-	-	+	30
+	-	+	20
-	+	+	10
+	+	+	25

Minitab Output

Binary Logistic Regression: Number Survi, Freq versus Sodium Citra, Glycerol, .

```
Link Function: Logit

Response Information

Variable         Value    Count
Number Survived  Success    168
                 Failure    232
Freq             Total      400

Logistic Regression Table
                                                         Odds     95% CI
Predictor             Coef   SE Coef      Z      P  Ratio  Lower  Upper
Constant         -0.376962  0.110113  -3.42  0.001
Sodium Citrate   0.0932642  0.110103   0.85  0.397   1.10   0.88   1.36
Glycerol         -0.463247  0.110078  -4.21  0.000   0.63   0.51   0.78
Equilbrium Time  0.0259045  0.109167   0.24  0.812   1.03   0.83   1.27
AB                0.585116  0.110066   5.32  0.000   1.80   1.45   2.23
AC               0.0543714  0.109317   0.50  0.619   1.06   0.85   1.31
BC                0.112190  0.108845   1.03  0.303   1.12   0.90   1.38

Log-Likelihood = -248.028
Test that all slopes are zero: G = 48.178, DF = 6, P-Value = 0.000

Goodness-of-Fit Tests

Method           Chi-Square  DF      P
Pearson            0.113790   1  0.736
Deviance           0.113865   1  0.736
Hosmer-Lemeshow    0.113790   6  1.000
```

This analysis shows that Glycerol (*B*) and the Sodium Citrate x Glycerol (*AB*) interaction have an effect on the survival rate of spermatozoa.

15.11. A soft drink distributor is studying the effectiveness of delivery methods. Three different types of hand trucks have been developed, and an experiment is performed in the company's methods engineering laboratory. The variable of interest is the delivery time in minutes (y); however, delivery time is also strongly related to the case volume delivered (x). Each hand truck is used four times and the data that follow are obtained. Analyze the data and draw the appropriate conclusions. Use α=0.05.

		Hand	Truck	Type	
1	1	2	2	3	3
y	*x*	*y*	*x*	*y*	*x*
27	24	25	26	40	38
44	40	35	32	22	26
33	35	46	42	53	50
41	40	26	25	18	20

From the analysis performed in *Minitab*, hand truck does not have a statistically significant effect on delivery time. Volume, as expected, does have a significant effect.

Minitab Output

General Linear Model: Time versus Truck

```
Factor     Type Levels Values
Truck     fixed      3 1 2 3

Analysis of Variance for Time, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Volume      1    1232.07    1217.55    1217.55  232.20  0.000
Truck       2      11.65      11.65       5.82    1.11  0.375
Error       8      41.95      41.95       5.24
Total      11    1285.67

Term          Coef   SE Coef        T      P
Constant    -4.747     2.638    -1.80  0.110
Volume     1.17326   0.07699    15.24  0.000
```

15.12. Compute the adjusted treatment means and the standard errors of the adjusted treatment means for the data in Problem 15.11.

$$\text{adj } \bar{y}_{i.} = \bar{y}_{i.} - \ddot{\beta}(\bar{x}_{i.} - \bar{x}_{..})$$

$$\text{adj } \bar{y}_{1.} = \frac{145}{4} - (1.173)\left(\frac{139}{4} - \frac{398}{12}\right) = 34.39$$

$$\text{adj } \bar{y}_{2.} = \frac{132}{4} - (1.173)\left(\frac{125}{4} - \frac{398}{12}\right) = 35.25$$

$$\text{adj } \bar{y}_{3.} = \frac{133}{4} - (1.173)\left(\frac{134}{4} - \frac{398}{12}\right) = 32.86$$

$$S_{adj.\bar{y}_{i.}} = \left[MS_E\left\{\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}}\right\}\right]^{\frac{1}{2}}$$

$$S_{adj.\bar{y}_{1.}} = \left[5.24\left\{\frac{1}{4} + \frac{(34.75 - 33.17)^2}{884.50}\right\}\right]^{\frac{1}{2}} = 1.151$$

$$S_{adj.\bar{y}_{2.}} = \left[5.24\left\{\frac{1}{4}+\frac{(31.25-33.17)^2}{884.50}\right\}\right]^{\frac{1}{2}} = 1.154$$

$$S_{adj.\bar{y}_{3.}} = \left[5.24\left\{\frac{1}{4}+\frac{(33.50-33.17)^2}{884.50}\right\}\right]^{\frac{1}{2}} = 1.145$$

The solutions can also be obtained with *Minitab* as follows:

Minitab Output

```
Least Squares Means for Time

Truck      Mean   SE Mean
1         34.39     1.151
2         35.25     1.154
3         32.86     1.145
```

15.13. The sums of squares and products for a single-factor analysis of covariance follow. Complete the analysis and draw appropriate conclusions. Use $\alpha = 0.05$.

Source of Variation	Degrees of Freedom	Sums of Squares and Products *x*	*xy*	*y*
Treatment	3	1500	1000	650
Error	12	6000	1200	550
Total	15	7500	2200	1200

	Sums of Squares and Products				Adjusted			
Source	*df*	*x*	*xy*	*y*	*y*	*df*	*MS*	F_0
Treatment	3	1500	1000	650	-	-		
Error	12	6000	1200	550	310	11	28.18	
Total	15	7500	2200	1200	554.67	14		
Adjusted	Treat.				244.67	3	81.56	2.89

Treatments differ only at 10%.

15.14. Find the standard errors of the adjusted treatment means in Example 15.5.

From Example 15.5, $\bar{y}_{1.} = 40.38$, adj $\bar{y}_{2.} = 41.42$, adj $\bar{y}_{3.} = 38.80$

$$S_{adj.\bar{y}_{1.}} = \left[2.54\left\{\frac{1}{5}+\frac{(25.20-24.13)^2}{195.60}\right\}\right]^{\frac{1}{2}} = 0.7231$$

$$S_{adj.\bar{y}_{2.}} = \left[2.54\left\{\frac{1}{5}+\frac{(26.00-24.13)^2}{195.60}\right\}\right]^{\frac{1}{2}} = 0.7439$$

$$S_{adj.\bar{y}_{3.}} = \left[2.54\left\{\frac{1}{5}+\frac{(21.20-24.13)^2}{195.60}\right\}\right]^{\frac{1}{2}} = 0.7871$$

15.15. Four different formulations of an industrial glue are being tested. The tensile strength of the glue when it is applied to join parts is also related to the application thickness. Five observations on strength (y) in pounds and thickness (x) in 0.01 inches are obtained for each formulation. The data are shown in the following table. Analyze these data and draw appropriate conclusions.

			Glue	Formulation			
1	1	2	2	3	3	4	4
y	*x*	*y*	*x*	*y*	*x*	*y*	*x*
46.5	13	48.7	12	46.3	15	44.7	16
45.9	14	49.0	10	47.1	14	43.0	15
49.8	12	50.1	11	48.9	11	51.0	10
46.1	12	48.5	12	48.2	11	48.1	12
44.3	14	45.2	14	50.3	10	48.6	11

From the analysis performed in *Minitab*, glue formulation does not have a statistically significant effect on strength. As expected, glue thickness does affect strength.

Minitab Output

General Linear Model: Strength versus Glue

```
Factor     Type Levels Values
Glue      fixed      4 1 2 3 4

Analysis of Variance for Strength, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Thick       1     68.852     59.566     59.566   42.62  0.000
Glue        3      1.771      1.771      0.590    0.42  0.740
Error      15     20.962     20.962      1.397
Total      19     91.585

Term          Coef   SE Coef        T      P
Constant    60.089     1.944    30.91  0.000
Thick      -1.0099    0.1547    -6.53  0.000

Unusual Observations for Strength

Obs  Strength       Fit      SE Fit  Residual   St Resid
  3   49.8000   47.5299      0.5508    2.2701      2.17R

R denotes an observation with a large standardized residual.

Expected Mean Squares, using Adjusted SS

Source       Expected Mean Square for Each Term
 1 Thick     (3) + Q[1]
 2 Glue      (3) + Q[2]
 3 Error     (3)

Error Terms for Tests, using Adjusted SS

Source       Error DF  Error MS  Synthesis of Error MS
 1 Thick        15.00     1.397  (3)
 2 Glue         15.00     1.397  (3)

Variance Components, using Adjusted SS

Source    Estimated Value
Error               1.397
```

15.16. Compute the adjusted treatment means and their standard errors using the data in Problem 15.15.

$$\text{adj } \bar{y}_{i.} = \bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})$$
$$\text{adj } \bar{y}_{1.} = 46.52 - (-1.0099)(13.00 - 12.45) = 47.08$$
$$\text{adj } \bar{y}_{2.} = 48.30 - (-1.0099)(11.80 - 12.45) = 47.64$$
$$\text{adj } \bar{y}_{3.} = 48.16 - (-1.0099)(12.20 - 12.45) = 47.91$$
$$\text{adj } \bar{y}_{4.} = 47.08 - (-1.0099)(12.80 - 12.45) = 47.43$$

$$S_{adj.\bar{y}_{i.}} = \left[MS_E \left\{ \frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}} \right\} \right]^{\frac{1}{2}}$$

$$S_{adj.\bar{y}_{1.}} = \left[1.40 \left\{ \frac{1}{5} + \frac{(13.00 - 12.45)^2}{58.40} \right\} \right]^{\frac{1}{2}} = 0.5360$$

$$S_{adj.\bar{y}_{2.}} = \left[1.40 \left\{ \frac{1}{5} + \frac{(11.80 - 12.45)^2}{58.40} \right\} \right]^{\frac{1}{2}} = 0.5386$$

$$S_{adj.\bar{y}_{3.}} = \left[1.40 \left\{ \frac{1}{5} + \frac{(12.20 - 12.45)^2}{58.40} \right\} \right]^{\frac{1}{2}} = 0.5306$$

$$S_{adj.\bar{y}_{4.}} = \left[1.40 \left\{ \frac{1}{5} + \frac{(12.80 - 12.45)^2}{58.40} \right\} \right]^{\frac{1}{2}} = 0.5319$$

The adjusted treatment means can also be generated in *Minitab* as follows:

Minitab Output
```
Least Squares Means for Strength

Glue      Mean   SE Mean
1        47.08    0.5355
2        47.64    0.5382
3        47.91    0.5301
4        47.43    0.5314
```

15.17. An engineer is studying the effect of cutting speed on the rate of metal removal in a machining operation. However, the rate of metal removal is also related to the hardness of the test specimen. Five observations are taken at each cutting speed. The amount of metal removed (y) and the hardness of the specimen (x) are shown in the following table. Analyze the data using and analysis of covariance. Use α=0.05.

		Cutting	Speed	(rpm)	
1000	1000	1200	1200	1400	1400
y	x	y	x	y	x
68	120	112	165	118	175
90	140	94	140	82	132
98	150	65	120	73	124
77	125	74	125	92	141
88	136	85	133	80	130

As shown in the analysis performed in *Minitab*, there is no difference in the rate of removal between the three cutting speeds. As expected, the hardness does have an impact on rate of removal.

Minitab Output

```
General Linear Model: Removal versus Speed

Factor     Type Levels Values
Speed     fixed      3 1000 1200 1400

Analysis of Variance for Removal, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
Hardness    1     3075.7     3019.3     3019.3  347.96  0.000
Speed       2        2.4        2.4        1.2    0.14  0.872
Error      11       95.5       95.5        8.7
Total      14     3173.6

Term          Coef   SE Coef        T      P
Constant   -41.656     6.907    -6.03  0.000
Hardness   0.93426   0.05008    18.65  0.000
Speed
1000         0.478     1.085     0.44  0.668
1200         0.036     1.076     0.03  0.974

Unusual Observations for Removal

Obs   Removal       Fit      SE Fit  Residual   St Resid
  8    65.000    70.491       1.558    -5.491     -2.20R

R denotes an observation with a large standardized residual.

Expected Mean Squares, using Adjusted SS

Source       Expected Mean Square for Each Term
 1 Hardness  (3) + Q[1]
 2 Speed     (3) + Q[2]
 3 Error     (3)

Error Terms for Tests, using Adjusted SS

Source       Error DF  Error MS  Synthesis of Error MS
 1 Hardness     11.00       8.7  (3)
 2 Speed        11.00       8.7  (3)

Variance Components, using Adjusted SS

Source    Estimated Value
Error               8.677

Means for Covariates

Covariate      Mean     StDev
Hardness      137.1     15.94

Least Squares Means for Removal

Speed      Mean   SE Mean
1000      86.88     1.325
1200      86.44     1.318
1400      85.89     1.328
```

15.18. Show that in a single factor analysis of covariance with a single covariate a $100(1-\alpha)$ percent confidence interval on the i_{th} adjusted treatment mean is

$$\bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \pm t_{\alpha/2,a(n-1)-1}\left[MS_E\left(\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}}\right)\right]^{\frac{1}{2}}$$

Using this formula, calculate a 95 percent confidence interval on the adjusted mean of machine 1 in Example 15.5.

The 100(1-α) percent interval on the i_{th} adjusted treatment mean would be

$$\bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \pm t_{\alpha/2,a(n-1)-1} S_{adj\bar{y}_{i.}}$$

where $\bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})$ is an estimator of the i_{th} adjusted treatment mean. The standard error of the adjusted treatment mean is found as follows:

$$V(adj.\bar{y}_{i.}) = V\left[\bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..})\right] = V(\bar{y}_{i.}) + (\bar{x}_{i.} - \bar{x}_{..})^2 V(\hat{\beta})$$

Since the $\{\bar{y}_{i.}\}$ and $\hat{\beta}$ are independent. From regression analysis, we have $V(\hat{\beta}) = \frac{\sigma^2}{E_{xx}}$. Therefore,

$$V(adj.\bar{y}_{i.}) = \frac{\sigma^2}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2 \sigma^2}{E_{xx}} = \sigma^2\left[\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}}\right]$$

Replacing σ^2 by its estimator MS_E, yields

$$\hat{V}(adj.\bar{y}_{i.}) = MS_E\left[\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}}\right] \text{ or}$$

$$S(adj.\bar{y}_{i.}) = \left\{MS_E\left[\frac{1}{n} + \frac{(\bar{x}_{i.} - \bar{x}_{..})^2}{E_{xx}}\right]\right\}^{\frac{1}{2}}$$

Substitution of this result into $\bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) \pm t_{\alpha/2,a(n-1)-1} S_{adj\bar{y}_{i.}}$ will produce the desired confidence interval. A 95% confidence interval on the mean of machine 1 would be found as follows:

$$adj.\bar{y}_{i.} = \bar{y}_{i.} - \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) = 40.38$$
$$S(adj.\bar{y}_{i.}) = 0.7231$$
$$[40.38 \pm t_{0.025,11}(0.7231)]$$
$$[40.38 \pm (2.20)(0.7231)]$$
$$[40.38 \pm 1.59]$$

Therefore, $38.79 \le \mu_1 \le 41.97$, where μ_1 denotes the true adjusted mean of treatment one.

15.19. Show that in a single-factor analysis of covariance with a single covariate, the standard error of the difference between any two adjusted treatment means is

$$S_{Adj\bar{y}_{i.}-Adj\bar{y}_{j.}}=\left[MS_E\left(\frac{2}{n}+\frac{\left(\bar{x}_{i.}-\bar{x}_{j.}\right)^2}{E_{xx}}\right)\right]^{\frac{1}{2}}$$

$$adj.\bar{y}_{i.}-adj.\bar{y}_{j.}=\bar{y}_{i.}-\hat{\beta}\left(\bar{x}_{i.}-\bar{x}_{..}\right)-\left[\bar{y}_{j.}-\hat{\beta}\left(\bar{x}_{j.}-\bar{x}_{..}\right)\right]$$

$$adj.\bar{y}_{i.}-adj.\bar{y}_{j.}=\bar{y}_{i.}-\bar{y}_{j.}-\hat{\beta}\left(\bar{x}_{i.}-\bar{x}_{j.}\right)$$

The variance of this statistic is

$$V\left[\bar{y}_{i.}-\bar{y}_{j.}-\hat{\beta}\left(\bar{x}_{i.}-\bar{x}_{j.}\right)\right]=V\left(\bar{y}_{i.}\right)+V\left(\bar{y}_{j.}\right)+\left(\bar{x}_{i.}-\bar{x}_{j.}\right)^2V\left(\hat{\beta}\right)$$
$$=\frac{\sigma^2}{n}+\frac{\sigma^2}{n}+\frac{\left(\bar{x}_{i.}-\bar{x}_{j.}\right)^2\sigma^2}{E_{xx}}=\sigma^2\left[\frac{2}{n}+\frac{\left(\bar{x}_{i.}-\bar{x}_{j.}\right)^2}{E_{xx}}\right]$$

Replacing σ^2 by its estimator MS_E, , and taking the square root yields the standard error

$$S_{Adj\bar{y}_{i.}-Adj\bar{y}_{j.}}=\left[MS_E\left(\frac{2}{n}+\frac{\left(\bar{x}_{i.}-\bar{x}_{j.}\right)^2}{E_{xx}}\right)\right]^{\frac{1}{2}}$$

15.20. Discuss how the operating characteristic curves for the analysis of variance can be used in the analysis of covariance.

To use the operating characteristic curves, fixed effects case, we would use as the parameter Φ^2,

$$\Phi^2=\frac{a\sum\tau_i^2}{n\sigma^2}$$

The test has *a*-1 degrees of freedom in the numerator and *a*(*n*-1)-1 degrees of freedom in the denominator.

15.21. Three different Pinot Noir wines were evaluated by a panel of eight judges. The judges are considered a random panel of all possible judges. The wines are evaluated on a 100-point scale. The wines were presented in random order to each judge, and the following results were obtained.

	Wine		
Judge	1	2	3
1	85	88	93
2	90	89	94
3	88	90	98
4	91	93	96
5	92	92	95
6	89	90	95

7	90	91	97
8	91	89	98

Analyze the data from this experiment. Is there a difference in the wine quality? Analyze the residuals and comment on model adequacy.

This is a repeated measures problem where the three Pinot Noirs are the treatments and the random Judges are the subjects. As described in the textbook, the analysis for a single factor repeated measures is the same as the randomized complete block design, RCBD, where the Judges are the random blocks.

The analysis below identifies a significant difference in the wine quality.

The residual plots do not identify any concerns with model adequacy.

Design Expert Output

Response 1 Wine Quality
ANOVA for selected factorial model
Analysis of variance table [Classical sum of squares - Type II]

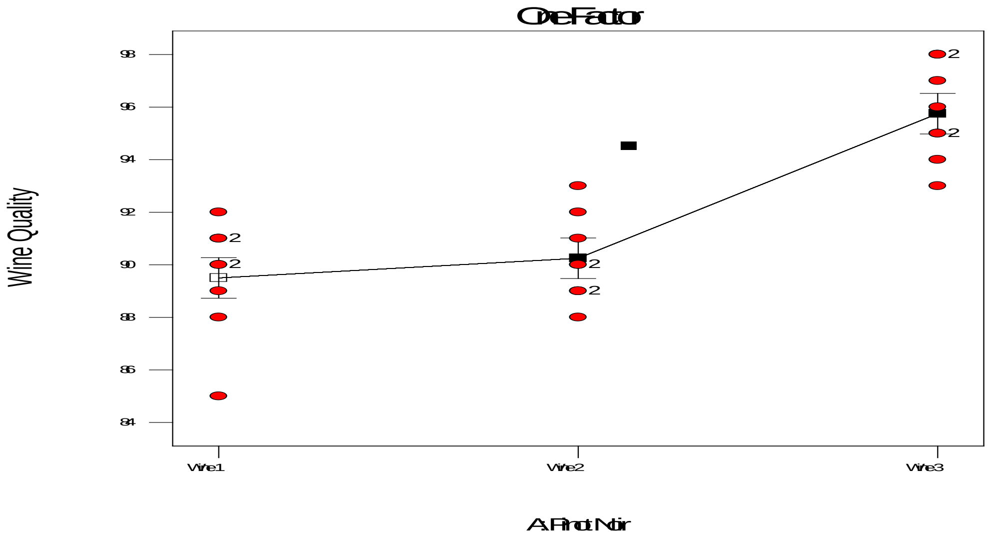
Source	**Sum of Squares**	**df**	**Mean Square**	**F Value**	**p-value Prob > F**	
Block	48.00	7	6.86			
Model	186.33	2	93.17	44.98	< 0.0001	significant
A-Pinot Noir	*186.33*	*2*	*93.17*	*44.98*	*< 0.0001*	
Residual	29.00	14	2.07			
Cor Total	263.33	23				

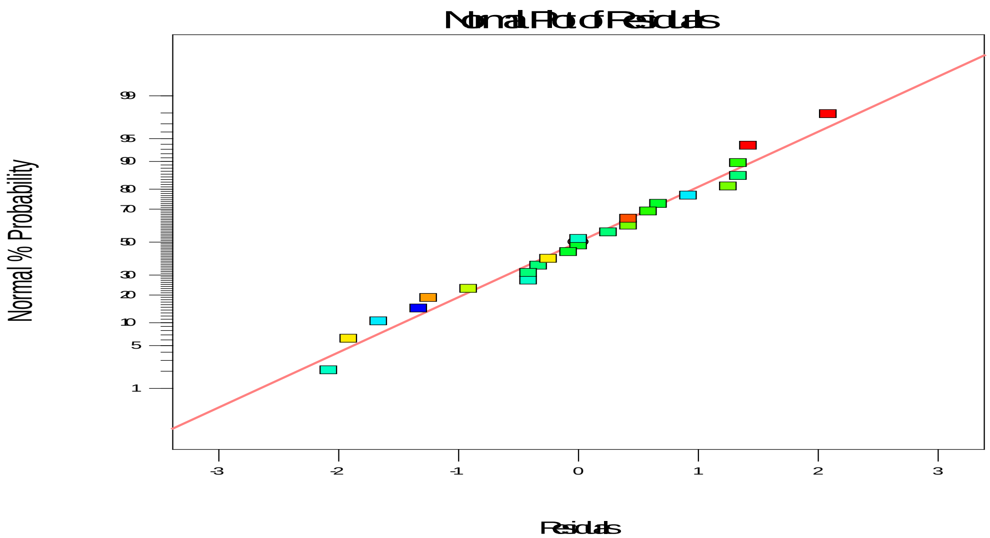
Std. Dev.	1.44	R-Squared	0.8653
Mean	91.83	Adj R-Squared	0.8461
C.V. %	1.57	Pred R-Squared	0.6042
PRESS	85.22	Adeq Precision	11.751

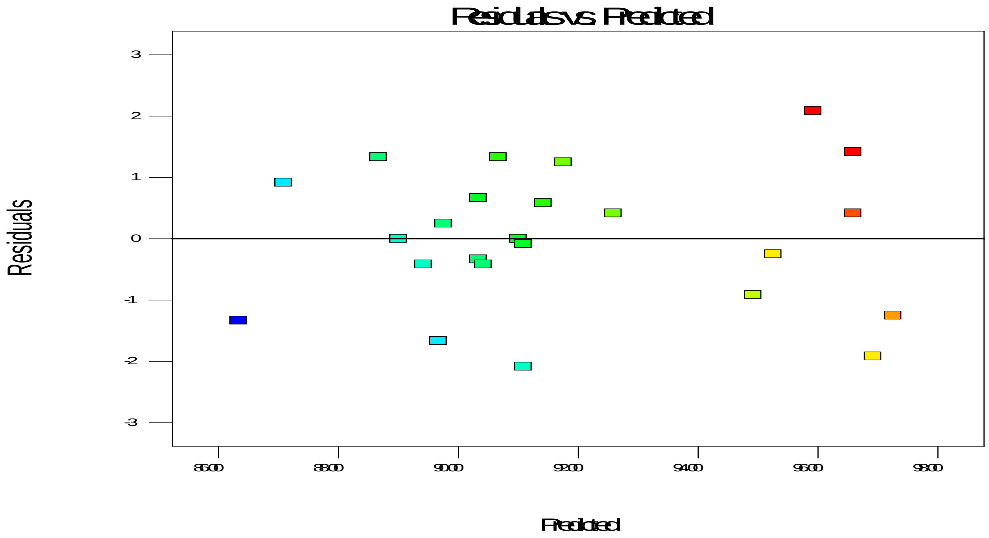
Treatment Means (Adjusted, If Necessary)

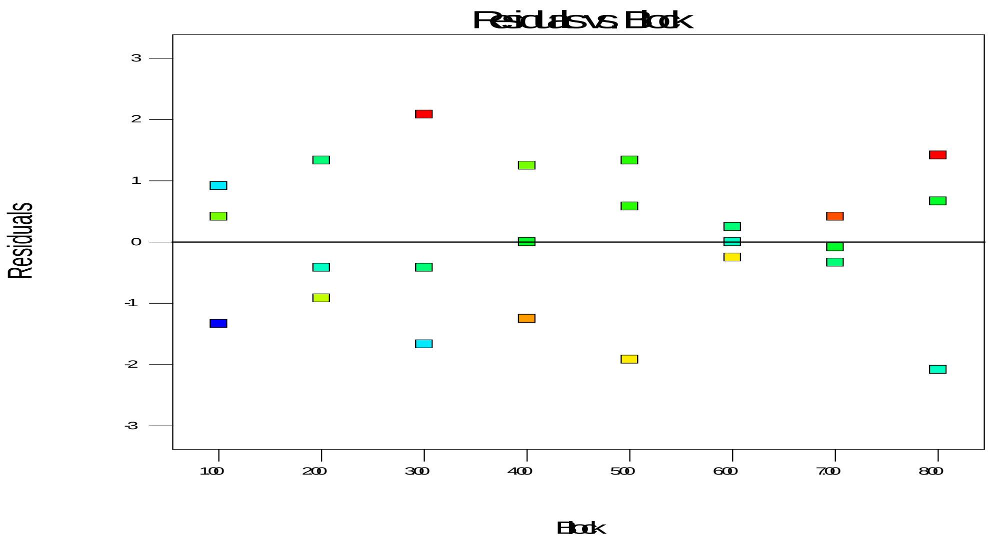
	Estimated Mean	**Standard Error**
1-Wine 1	89.50	0.51
2-Wine 2	90.25	0.51
3-Wine 3	95.75	0.51

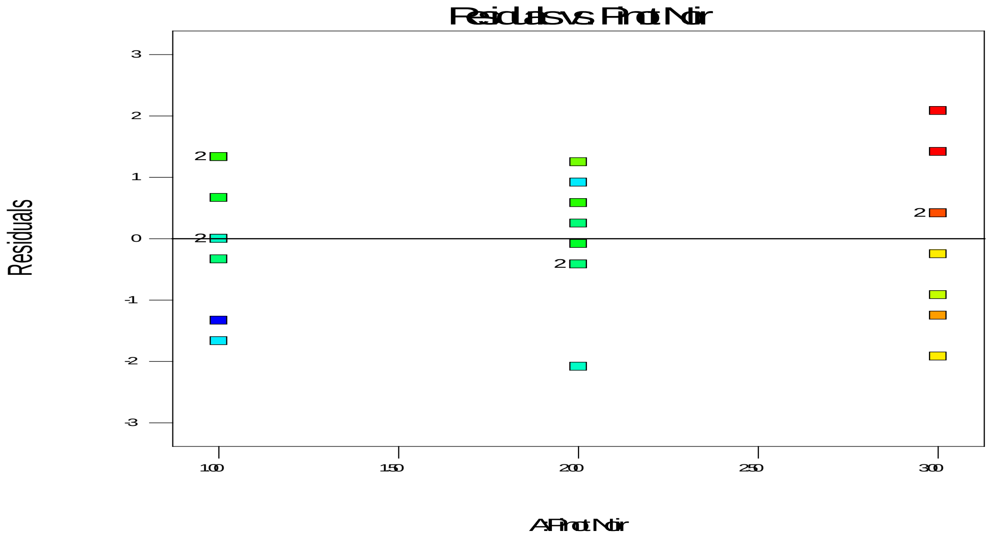
Treatment	**Mean Difference**	**df**	**Standard Error**	**t for H0 Coeff=0**	**Prob > |t|**
1 vs 2	-0.75	1	0.72	-1.04	0.3150
1 vs 3	-6.25	1	0.72	-8.69	< 0.0001
2 vs 3	-5.50	1	0.72	-7.64	< 0.0001


One Factor
Wine Quality
98
96
94
92
90
88
86
84
Wine1
Wine2
Wine3
A: Pinot Noir


Normal Plot of Residuals
Normal % Probability
99
95
90
80
70
50
30
20
10
5
1
-3
-2
-1
0
1
2
3
Residuals


Residuals vs Predicted
Residuals
3
2
1
0
-1
-2
-3
Predicted


Residuals vs Block
Residuals
3
2
1
0
-1
-2
-3
Block


Residuals vs Pinot Noir
Residuals
A:Pinot Noir
3
2
1
0
-1
-2
-3
100
150
200
250
300