STAT571B
Homework 5
Ken Youens-Clark

Montgomery 6.15

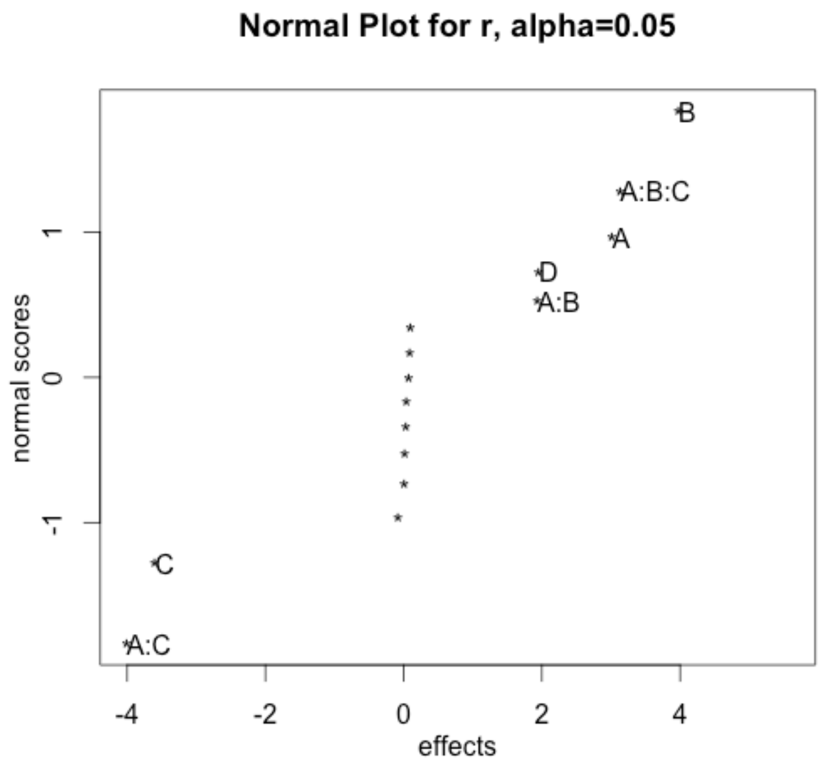
(a) Estimate the factor effects. Which factor effects appear to be large?

```
> vars = c("A","B","C","D")
> levels = c(-1, 1)
> reps = matrix(
+   c(7.037, 14.707, 11.635, 17.273, 10.403, 4.368, 9.360, 13.440,
+     8.561, 16.867, 13.876, 19.824, 11.846, 6.125, 11.190, 15.653,
+     6.376, 15.219, 12.089, 17.815, 10.151, 4.098, 9.253, 12.923,
+     8.951, 17.052, 13.658, 19.639, 12.337, 5.904, 10.935, 15.053),
+    ncol=2
+ )
> Total <- apply(reps,1,sum)
> exp = expand.grid(A=levels, B=levels, C=levels, D=levels)
>
> # create a data frame by binding the observations to the exp design
> dat = data.frame(cbind(exp, r=reps, Total))
>
> n = 2; A = dat$A; B = dat$B; C = dat$C; D = dat$D;
> AB = A*B; AC = A*C; AD = A*D; BC = B*C;
> BD = B*D; CD = C*D; ABC = A*B*C; ABD = A*B*D; BCD = B*C*D
> effects = data.frame(e=t(Total %*%
cbind(A,B,C,D,AB,AC,AD,BC,CD,ABC,ABD,BCD)/(8*n)))
> effects$half = effects$e/2
> effects
            e       half
A    3.018875  1.5094375
B    3.975875  1.9879375
C   -3.596250 -1.7981250
D    1.957750  0.9788750
AB   1.934125  0.9670625
AC  -4.007750 -2.0038750
AD   0.076500  0.0382500
BC   0.096000  0.0480000
CD  -0.076875 -0.0384375
ABC  3.137500  1.5687500
ABD  0.098000  0.0490000
BCD  0.035625  0.0178125
```

A, B, C, D, AB, AC, and ABC have the largest values. (As I will need half of the effect values for (c), I calculated them here.) A Daniel plot confirms the significance of these factors as only they are labeled:

```
> df = data.frame(rbind(cbind(exp, r=reps[,1]), cbind(exp, r=reps[,2])))
```

```
> for (v in vars ) {
+   df[[v]] = factor(df[[v]])
+ }
> lmod = lm(r ~ A * B * C * D, df)
> DanielPlot(lmod)
```



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

```
            Df Sum Sq Mean Sq  F value   Pr(>F)
A            1  72.91   72.91  898.339 1.74e-15
B            1 126.46  126.46 1558.172  < 2e-16
C            1 103.46  103.46 1274.822  < 2e-16
D            1  30.66   30.66  377.802 1.49e-12
A:B          1  29.93   29.93  368.739 1.79e-12
A:C          1 128.50  128.50 1583.256  < 2e-16
B:C          1   0.07    0.07    0.908    0.355
A:D          1   0.05    0.05    0.577    0.459
B:D          1   0.02    0.02    0.220    0.645
C:D          1   0.05    0.05    0.583    0.456
A:B:C        1  78.75   78.75  970.325 9.49e-16
A:B:D        1   0.08    0.08    0.947    0.345
A:C:D        1   0.00    0.00    0.036    0.852
B:C:D        1   0.01    0.01    0.125    0.728
A:B:C:D      1   0.00    0.00    0.020    0.890
Residuals   16   1.30    0.08
```

In support of the Daniel plot, the above ANOVA table shows that the effects from factors A, B, C, D, AB, AC, and ABC are all well below α=0.05, so we reject the null hypothesis and state that some factors have a significant effect on cracking.

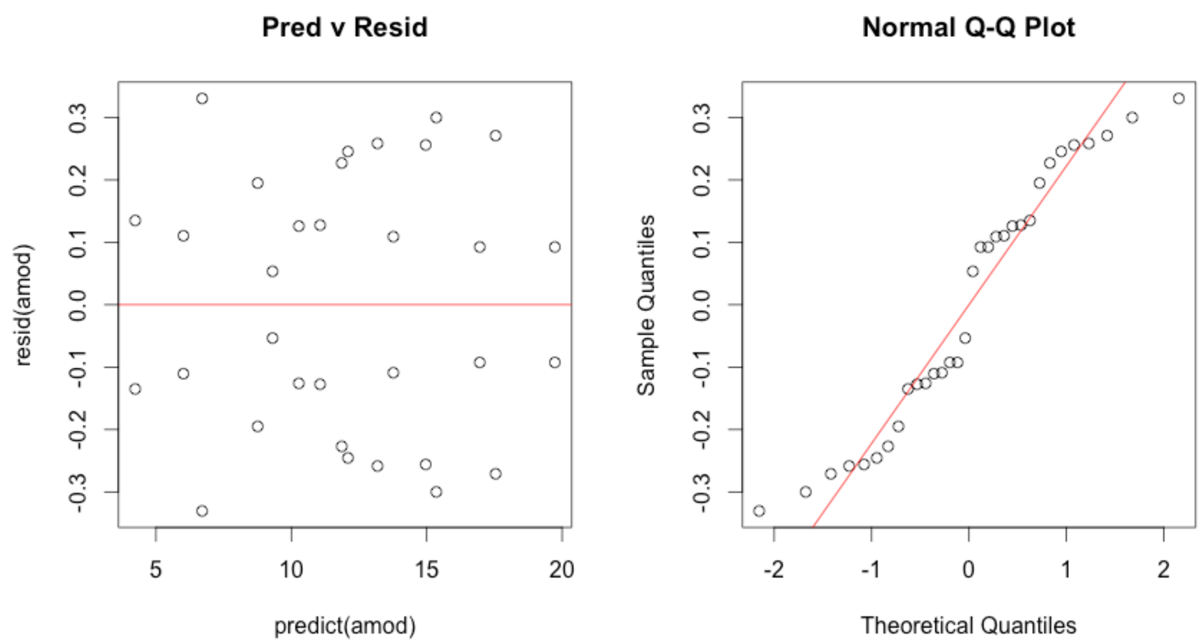
(c) Write down a regression model that can be used to predict crack length as a function of the significant main effects and interactions you have identified in part (b).

Using the average of the responses ("reps" above) as the intercept and 1/2 of the affect values from (a) for the most significant factors identified:

```
> sum(Total) / length(reps)
[1] 11.98806
y = 11.98806 + 1.509(A) + 1.988(B) + (-1.798)(C) + 0.979(D)
    + 0.967(AB) + (-2.004)(AC) + 1.569(ABC)
```

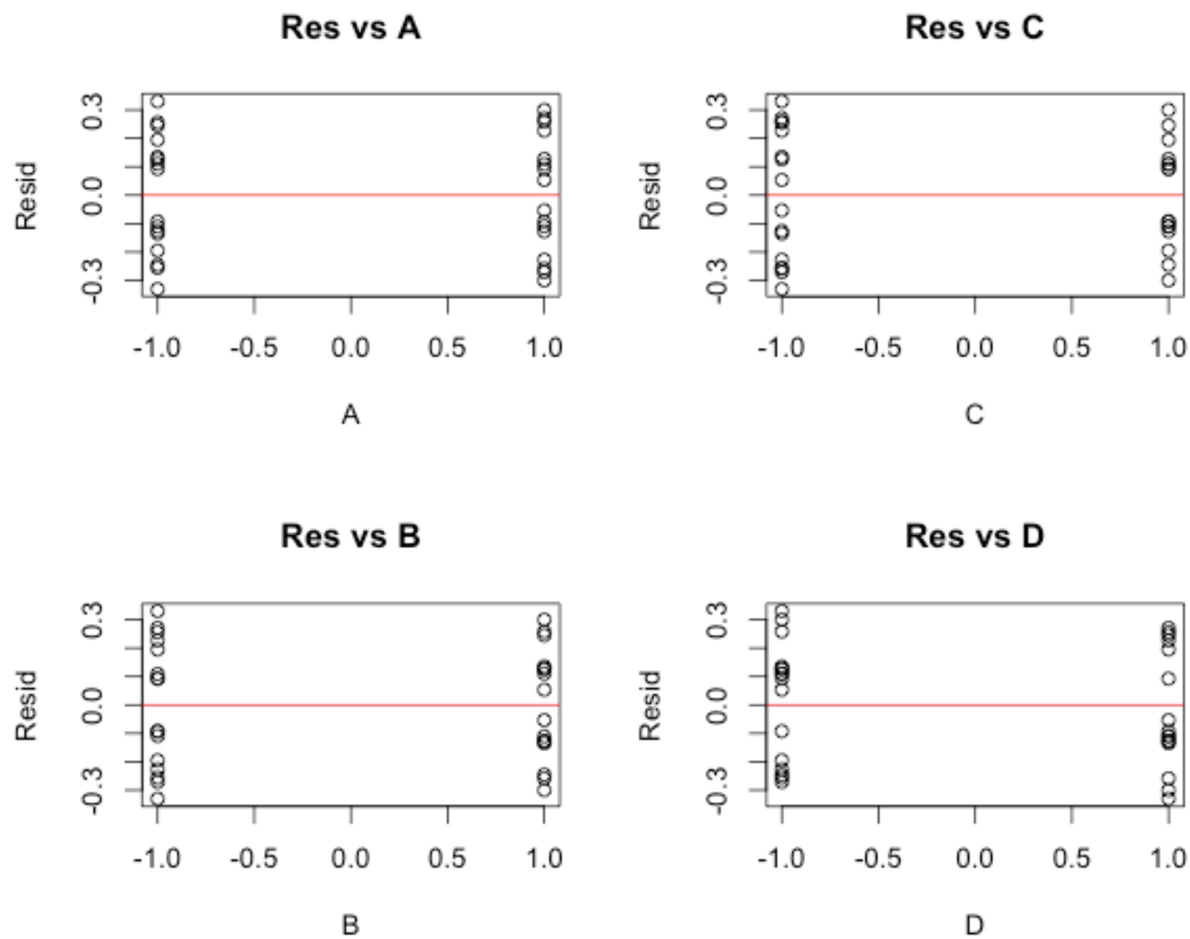
(d) Analyze the residuals from this experiment.

Inspection of the predicted vs. residuals shows an evenly and randomly distributed pattern, and the QQ plot of the residuals also shows fairly good fit.



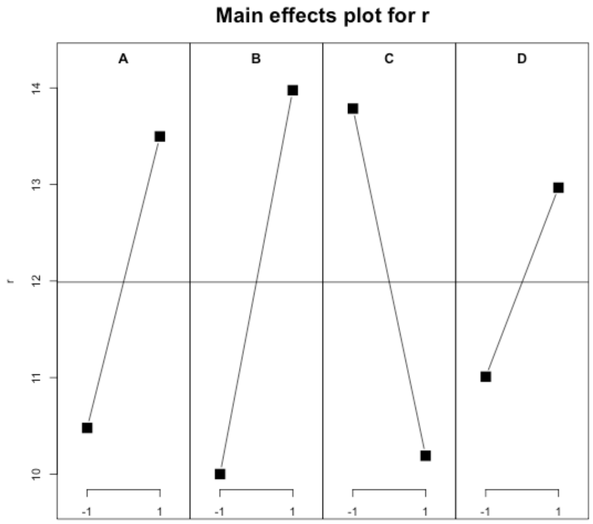
(e) Is there an indication that any of the factors affect the variability in cracking?

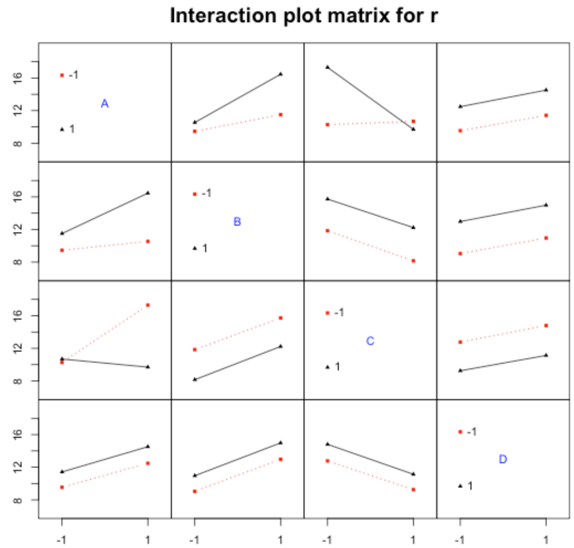
There really does not appear to be any patterns for each of the main factors vs the residuals, so I would say no.



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

Main effect and interaction plots:






```
A    3.018875
B    3.975875
C   -3.596250
D    1.957750 (from above)
```

Of the four factors pouring temperature (*A*), titanium content (*B*), heat treatment method (*C*), and amount of grain refiner used (*D*), higher values of A and B appear to cause more cracking. I would recommend using lower values of A and B and a high value of C. Factor D has no interaction with the others, so either high or low.

Montgomery 6.16

(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?

Using the above regression:

```
y = 11.98806 + 1.509(A) + 1.988(B) + (-1.798)(C) + 0.979(D)
    + 0.967(AB) + (-2.004)(AC)  + 1.569(ABC)
```

Given C = 1:

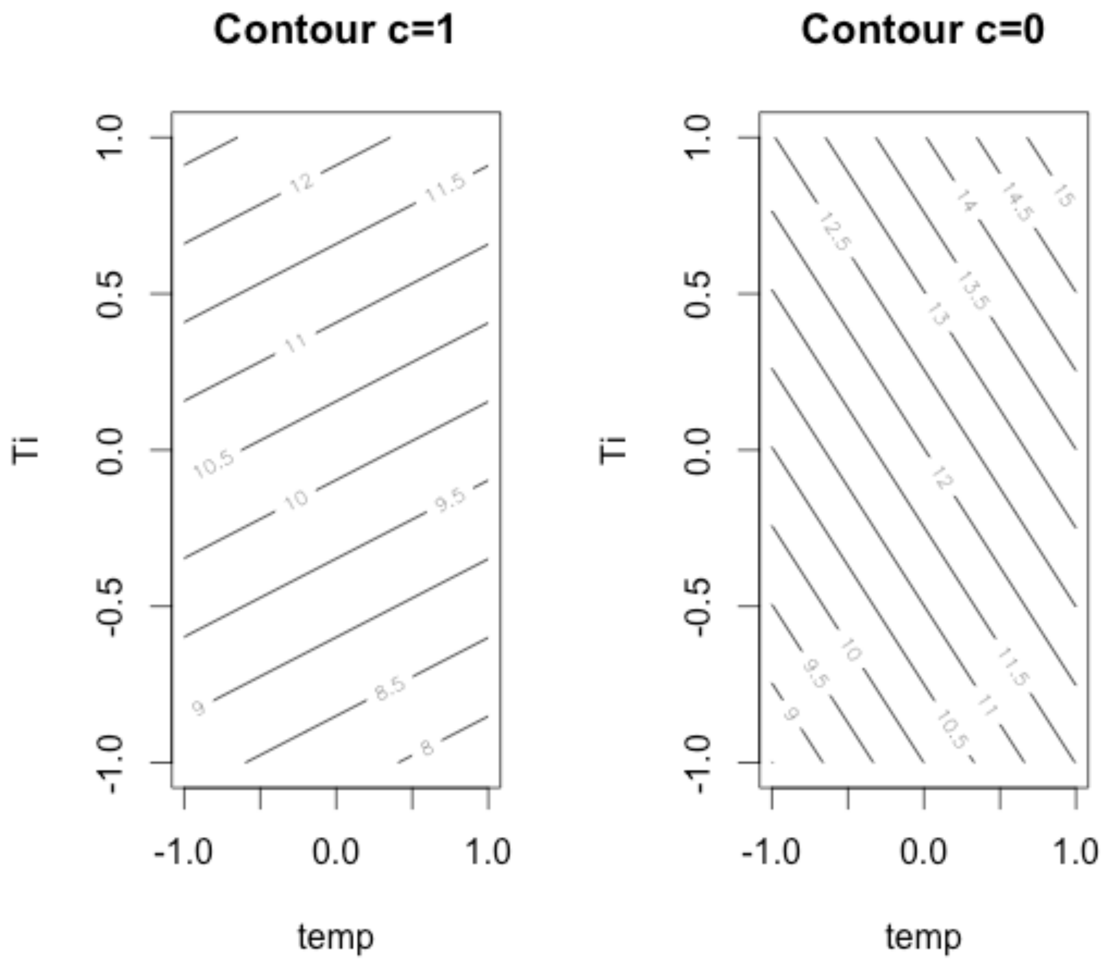
```
y = 11.98806 + 1.509(A) + 1.988(B) + (-1.798)(1) + 0.979(D)
    + 0.967(AB) + (-2.004)(A*1)  + 1.569(AB*1)
y = 10.19006 - 0.498(A) + 1.988(B) + 0.979(D) + 2.536(AB)
```

Given C = 0:

```
y = 11.98806 + 1.509(A) + 1.988(B) + (-1.798)(0) + 0.979(D)
    + 0.967(AB) + (-2.004)(A*0)  + 1.569(AB*0)
```

```
y = 11.98806 + 1.509(A) + 1.988(B) + 0.979(D) + 0.967(AB)
```

(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 = +?

Lower responses (less cracking) found with higher temperature and less titanium.

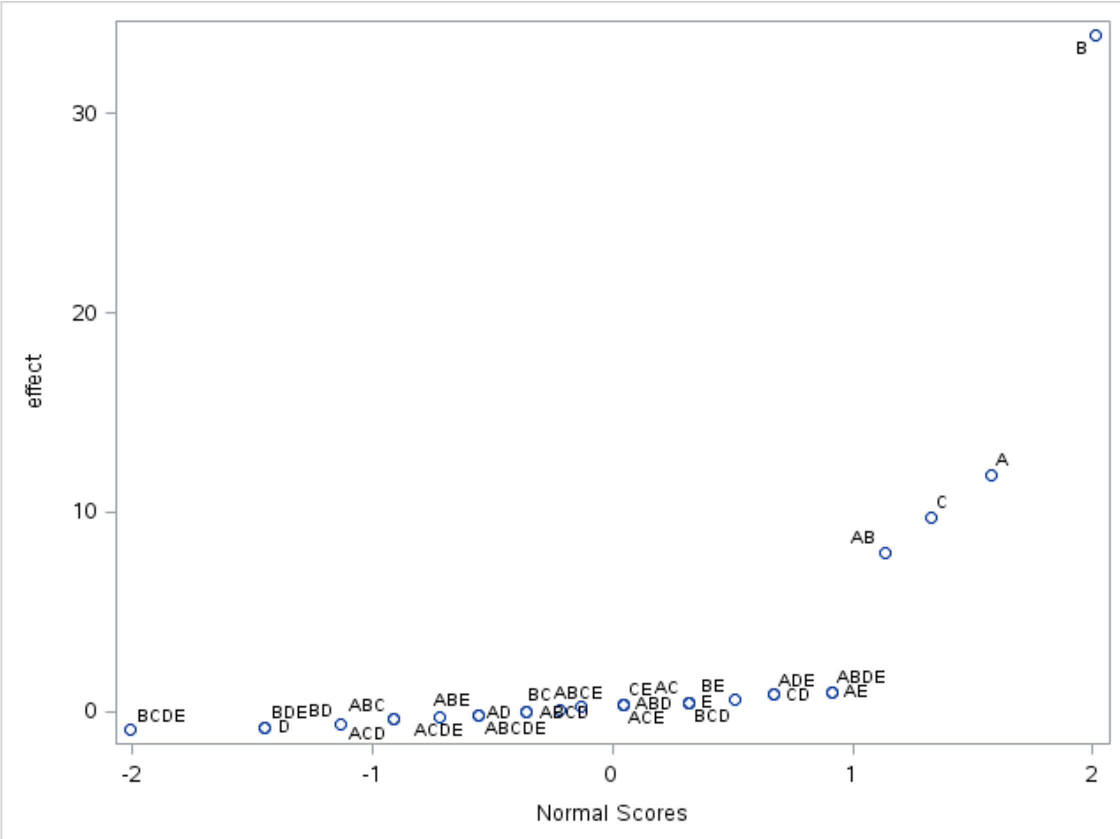
(d) Repeat part (c) assuming that you wish to use heat treatment method C = -.

I would recommend lower temperatures and less titanium.

Montgomery 6.26 (skip part h)

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

The effects from B, A, C, and AB appear to be significant.



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

Source	DF	Type III SS	Mean Square	F Value	Pr > F
A	1	1116.281250	1116.281250	382.27	<.0001
B	1	9214.031250	9214.031250	3155.34	<.0001
C	1	750.781250	750.781250	257.10	<.0001
AB	1	504.031250	504.031250	172.61	<.0001

The ANOVA table below show the p-values for A, B, C, and AB to be well below $\alpha=0.05$, therefore we reject the null hypothesis and state that they have a significant effect on the outcome.

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

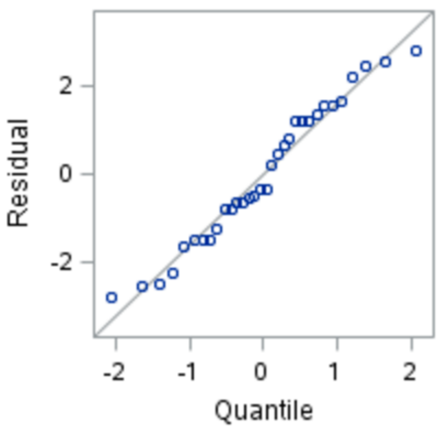
Using the values from the "parameter estimates" table:

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	30.53125	0.30208	101.07	<.0001
A	1	5.90625	0.30208	19.55	<.0001
B	1	16.96875	0.30208	56.17	<.0001
C	1	4.84375	0.30208	16.03	<.0001
AB	1	3.96875	0.30208	13.14	<.0001

```
y = 30.53125 + 5.90625(A) + 16.96875(B) + 4.84375(C) + 3.96875(AB)
```

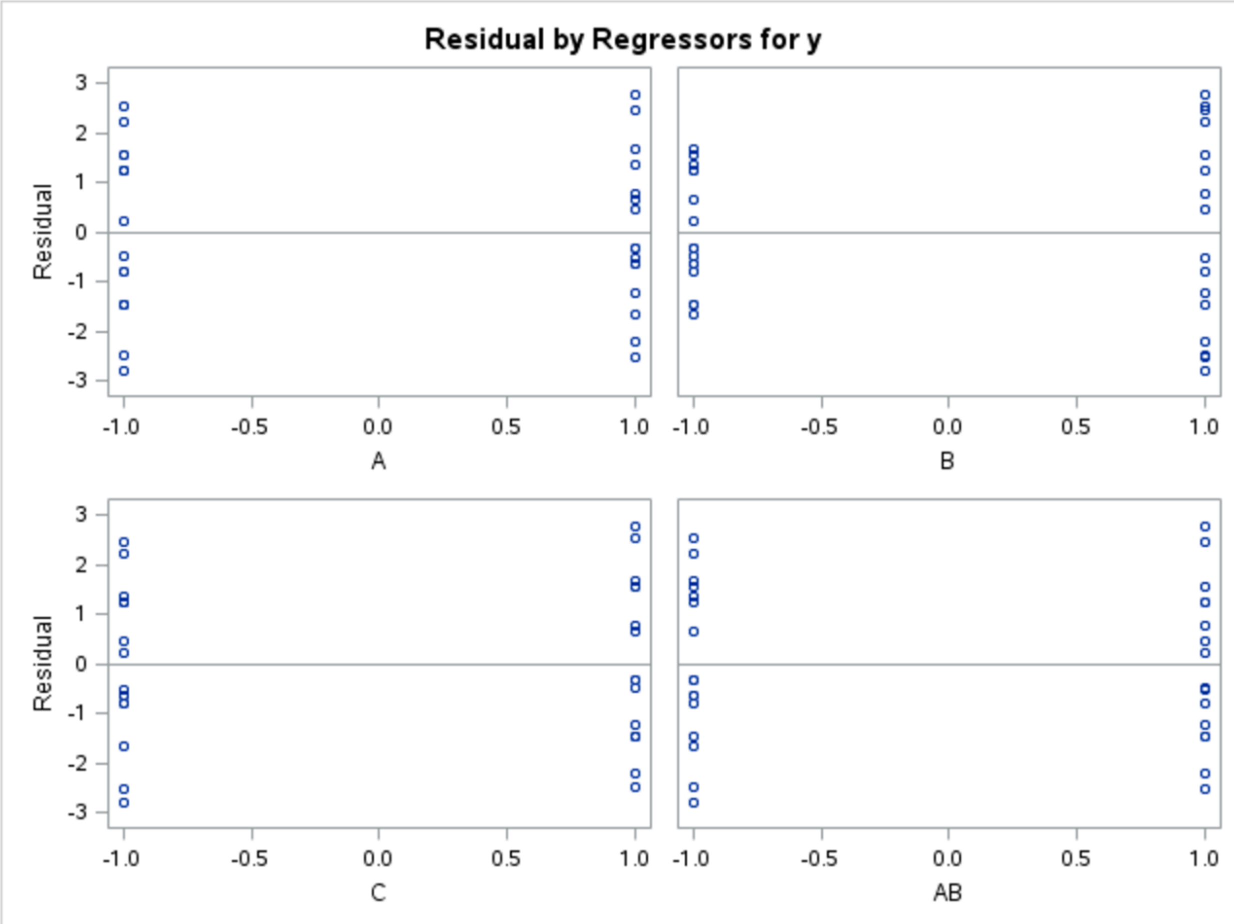
(d) Plot the residuals on normal probability paper. Is the plot satisfactory?

All points hew closely to the line, so the data looks normal.



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

I see no worrying trends in the plots for A, B, C, or AB.



(f) Interpret any significant interactions.

Unable to find how to create interaction plots in SAS.

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

Unable to make any without interaction plots.

Montgomery 7.2

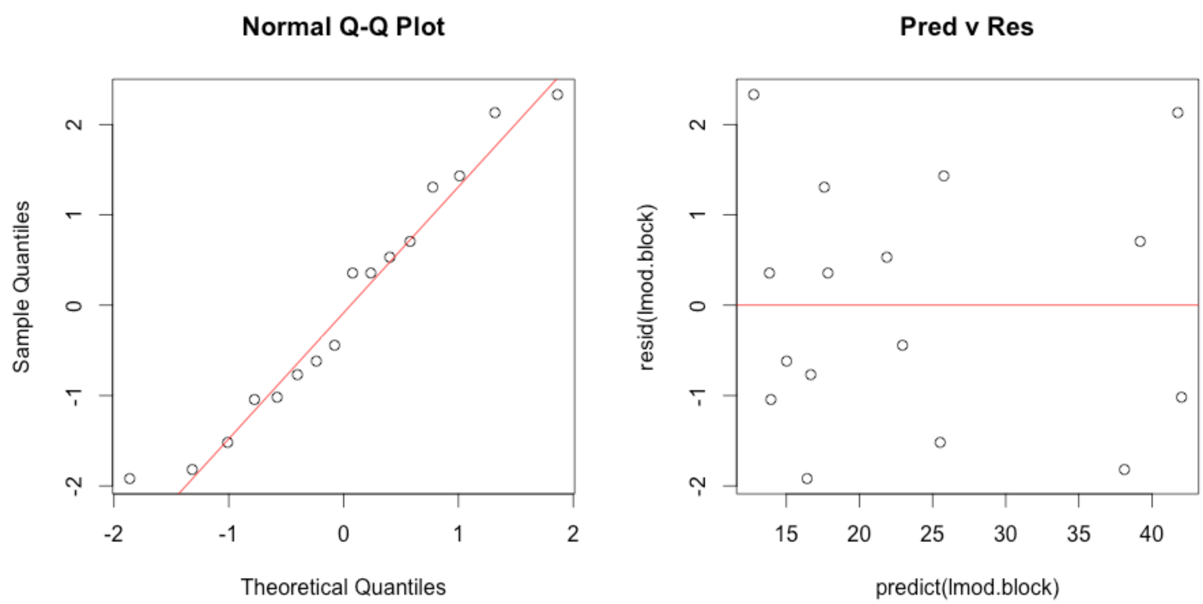
(a) Analyze the data from this experiment.

```
> reps = data.frame(
+   r1 = c(18.2, 27.2, 15.9, 41.0), r2 = c(18.9, 24.0, 14.5, 43.9),
+   r3 = c(12.9, 22.4, 15.1, 36.3), r4 = c(14.4, 22.5, 14.2, 39.9))
> levels = factor(c(-1, 1))
> exp = expand.grid(A=levels, B=levels)
> df = data.frame(rbind(
+   cbind(exp, r=reps$r1, block="I"),   cbind(exp, r=reps$r2, block="II"),
```

```
+   cbind(exp, r=reps$r3, block="III"), cbind(exp, r=reps$r4, block="IV")
+ ))
> options(show.signif.stars=FALSE)
> summary(aov(r ~ block + A * B, df))
            Df Sum Sq Mean Sq F value   Pr(>F)
block        3   44.4    14.8   4.864    0.028
A            1 1107.2  1107.2 364.211 1.37e-08
B            1  227.3   227.3  74.753 1.18e-05
A:B          1  303.6   303.6  99.876 3.60e-06
Residuals    9   27.4     3.0
```

The above ANOVA shows p-values for factors A (bit size) and B (drilling speed) to be well below α=0.05, so I would reject the null hypothesis and say that there is a significant difference when these are varied.

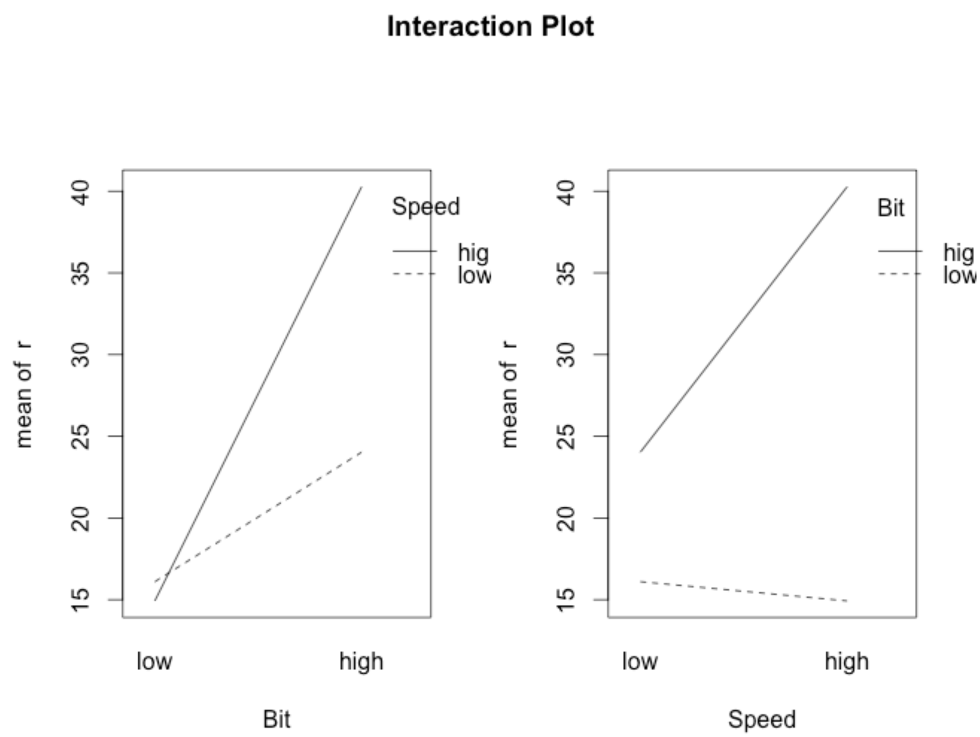
(b) Construct a normal probability plot of the residuals, and plot the residuals versus the predicted vibration level. Interpret these plots.

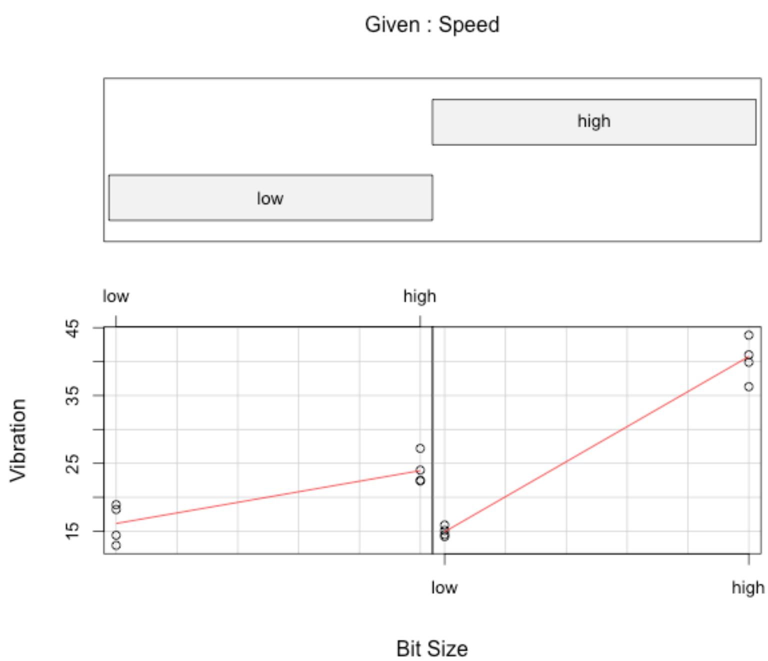


The QQ plots shows a good fit for with just one value (lower left) appearing to be an outlier. A scatterplot of predicted values vs residuals shows a random distribution evenly distributed above and below 0. Both of these support acceptance of normally distributed data.

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

Below are interaction and co-plots for bit size vs drilling speed:





When bit size is low and cutting speed is high, they interact to create lower vibration, so I would recommend to use the smaller 1/16 bit and higher 90 rpm cutting speed.