MIT OpenCourseWare
http://ocw.mit.edu

2.830J / 6.780J / ESD.63J Control of Manufacturing Processes (SMA 6303)
Spring 2008

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

HW 7 Solution 2008

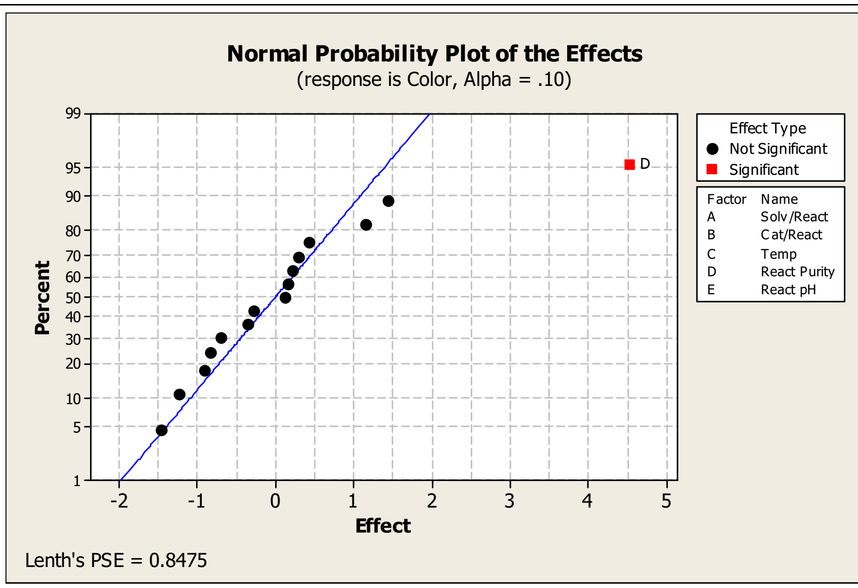
Problem 1 (12-10)

(a)

Factorial Fit: Color versus Solv/React, Cat/React, ...

Estimated Effects and Coefficients for Color (coded units)

Term	Effect	Coef
Constant		2.7700
Solv/React	1.4350	0.7175
Cat/React	-1.4650	-0.7325
Temp	-0.2725	-0.1363
React Purity	4.5450	2.2725
React pH	-0.7025	-0.3513
Solv/React*Cat/React	1.1500	0.5750
Solv/React*Temp	-0.9125	-0.4562
Solv/React*React Purity	-1.2300	-0.6150
Solv/React*React pH	0.4275	0.2138
Cat/React*Temp	0.2925	0.1462
Cat/React*React Purity	0.1200	0.0600
Cat/React*React pH	0.1625	0.0812
Temp*React Purity	-0.8375	-0.4187
Temp*React pH	-0.3650	-0.1825
React Purity*React pH	0.2125	0.1062

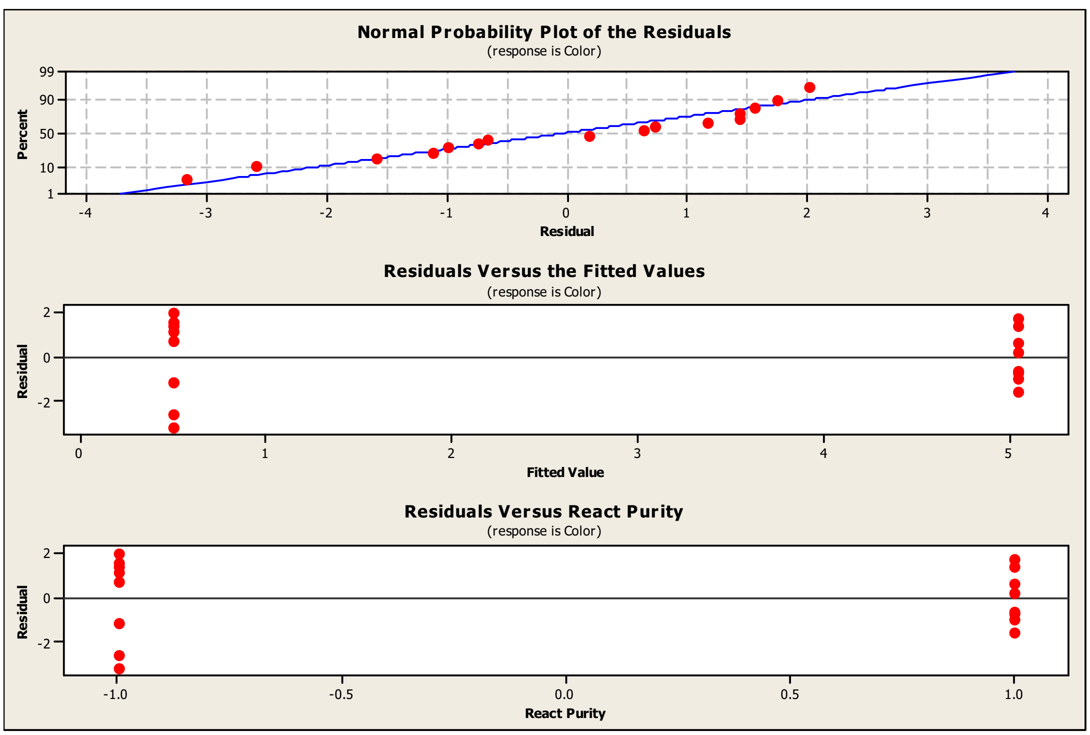


From visual examination of the normal probability plot of effects, only factor D (reactant purity) is significant.   Re-fit and analyze the reduced model.

Factorial Fit: Color versus React Purity

```
Estimated Effects and Coefficients for Color (coded units)

Term          Effect   Coef  SE Coef     T      P
Constant              2.770   0.4147  6.68  0.000
React Purity   4.545  2.272   0.4147  5.48  0.000


S = 1.65876   R-Sq = 68.20%   R-Sq(adj) = 65.93%


Analysis of Variance for Color (coded units)

Source          DF  Seq SS  Adj SS  Adj MS      F      P
Main Effects     1   82.63   82.63  82.628  30.03  0.000
Residual Error  14   38.52   38.52   2.751
  Pure Error    14   38.52   38.52   2.751
Total           15  121.15
```

(b)

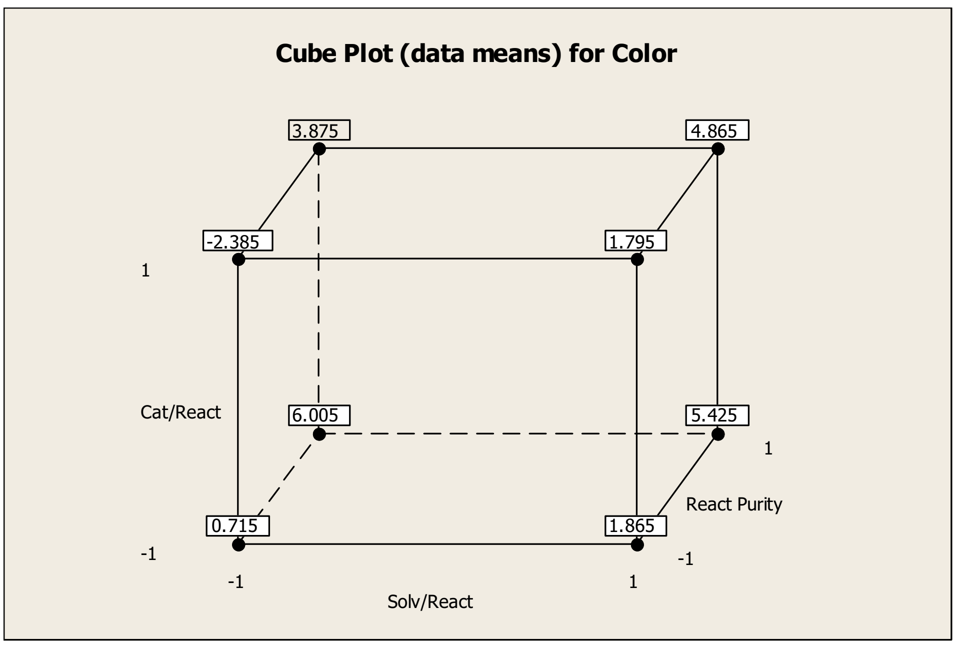


Residual plots indicate that there may be problems with both the normality and constant variance assumptions.

(c)

There is only one significant factor, D (reactant purity), so this design collapses to a one-factor experiment, or simply a 2-sample *t*-test.

Looking at the original normal probability plot of effects and effect estimates, the 2nd and 3rd largest effects in absolute magnitude are A (solvent/reactant) and B (catalyst/reactant). A cube plot in these factors shows how the design can be collapsed into a replicated 2^3 design. The highest color scores are at high reactant purity; the lowest at low reactant purity.

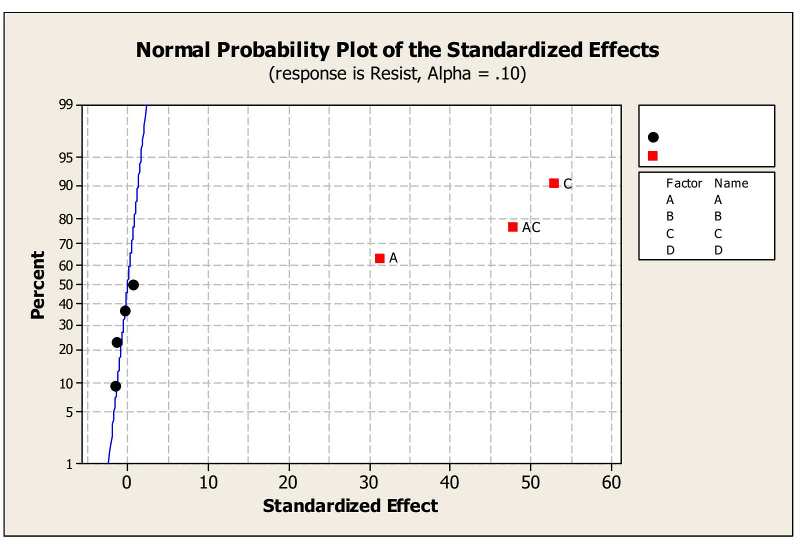


Problem 2 (12-15)

(a)

Factorial Fit: Resist versus A, B, C, D

```
Estimated Effects and Coefficients for Resist (coded units)

Term      Effect    Coef  SE Coef      T      P

Constant           60.433   0.6223  97.12  0.000

A         47.700  23.850   0.7621  31.29  0.000 *

B         -0.500  -0.250   0.7621  -0.33  0.759

C         80.600  40.300   0.7621  52.88  0.000 *

D         -2.400  -1.200   0.7621  -1.57  0.190

A*B        1.100   0.550   0.7621   0.72  0.510

A*C       72.800  36.400   0.7621  47.76  0.000 *

A*D       -2.000  -1.000   0.7621  -1.31  0.260

…

Analysis of Variance for Resist (coded units)

Source             DF   Seq SS   Adj SS   Adj MS      F      P

Main Effects        4  17555.3  17555.3  4388.83  944.51  0.000

2-Way Interactions  3  10610.1  10610.1  3536.70  761.13  0.000

Residual Error      4     18.6     18.6     4.65

  Curvature         1      5.6      5.6     5.61    1.30  0.338

  Pure Error        3     13.0     13.0     4.33

Total              11  28184.0
```



Examining the normal probability plot of effects, the main effects A and C and their two-factor interaction (AC) are significant. Re-fit and analyze a reduced model containing A, C, and AC.

(b)

Factorial Fit: Resist versus A, C

Estimated Effects and Coefficients for Resist (coded units)

Term	Effect	Coef	SE Coef	T	P	
Constant		60.43	0.6537	92.44	0.000	
A	47.70	23.85	0.8007	29.79	0.000	*
C	80.60	40.30	0.8007	50.33	0.000	*
A*C	72.80	36.40	0.8007	45.46	0.000	*

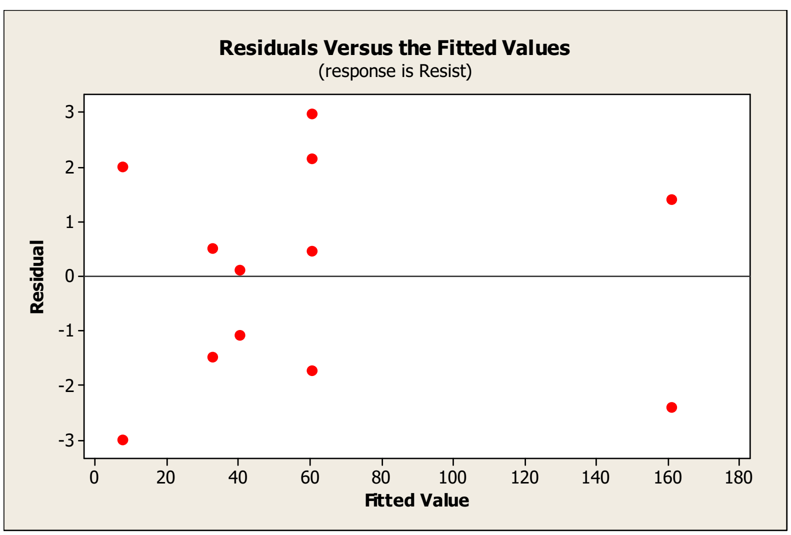
…

Analysis of Variance for Resist (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	2	17543.3	17543.3	8771.6	1710.43	0.000
2-Way Interactions	1	10599.7	10599.7	10599.7	2066.89	0.000
Residual Error	8	41.0	41.0	5.1		
Curvature	1	5.6	5.6	5.6	1.11	0.327
Pure Error	7	35.4	35.4	5.1		
Total	11	28184.0				

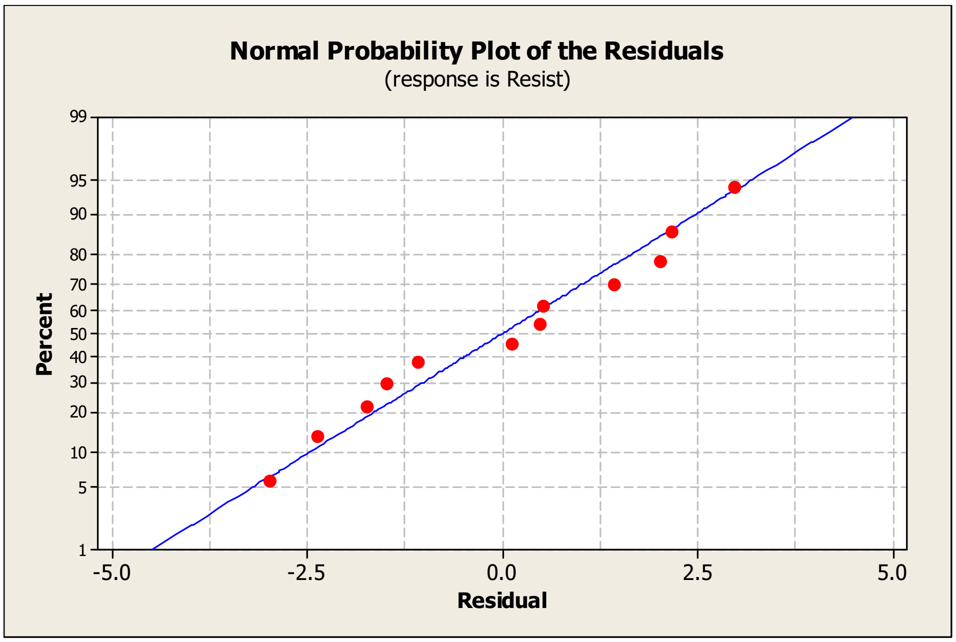
Curvature is not significant (P-value = 0.327), so continue with analysis.

(c)



A funnel pattern at the low value and an overall lack of consistent width suggest a problem with equal variance across the prediction range.

(d)



The normal probability plot of residuals is satisfactory.

The concern with variance in the predicted resistivity indicates that a data transformation may be needed.

Addendum to solution to Problem 2: manual test of curvature (courtesy R. Schwenke)

Test for curvature:

$$SS_{\text{pure quadratic}} = \frac{n_F \cdot n_C (\bar{y}_F - \bar{y}_C)^2}{n_F + n_C} \quad (dof = 1)$$

$$= \frac{8 \cdot 4 (59.95 - 61.4)^2}{8+4}$$

$$= 5.6067$$

$$\hat{\sigma}^2 = \frac{(63.4-61.4)^2 + (62.6-61.4)^2 + (58.7-61.4)^2 + (60.9-61.4)^2}{3}$$

$$= 4.3267$$

$$\Rightarrow F_0 = \frac{5.6067}{4.3267} = 1.2958$$

$$\Rightarrow P = \text{Fcdf}(1.2958) = 0.3377$$

with $\nu_1 = 1$
$\nu_2 = 3$

→ only 33.77% P-value for curvature. → no statistic evidence.

Problem 3: example solution (courtesy X. Su)

a.

Run		I	A	B	AB	Replicate Results					Totals	Effects estimate
			Design Factors									
1	(1)	1	-1	-1	1	0.1963	0.2185	0.1914	0.1814	0.2092	0.9968	
2	a	1	1	-1	-1	0.0914	0.0891	0.0925	0.0855	0.0913	0.4498	-0.07724
3	b	1	-1	1	-1	0.1107	0.1071	0.1109	0.1115	0.1145	0.5547	-0.05626
4	ab	1	1	1	1	0.065	0.065	0.0667	0.0662	0.0664	0.3293	0.03216

Source of variation	Sum of squares	Degrees of Freedom	Mean Square	F_0	P-value
A	0.02983	1	0.02983	628	1.40194E-16
B	0.015826	1	0.015826	333.1789	6.13524E-14
AB	0.005171	1	0.005171	108.8632	1.54196E-09
Curvature	0.000856	1	0.000856	18.02105	0.000396599
Residual Error	0.00095	20	4.75234E-05		
Total	0.052634	24			

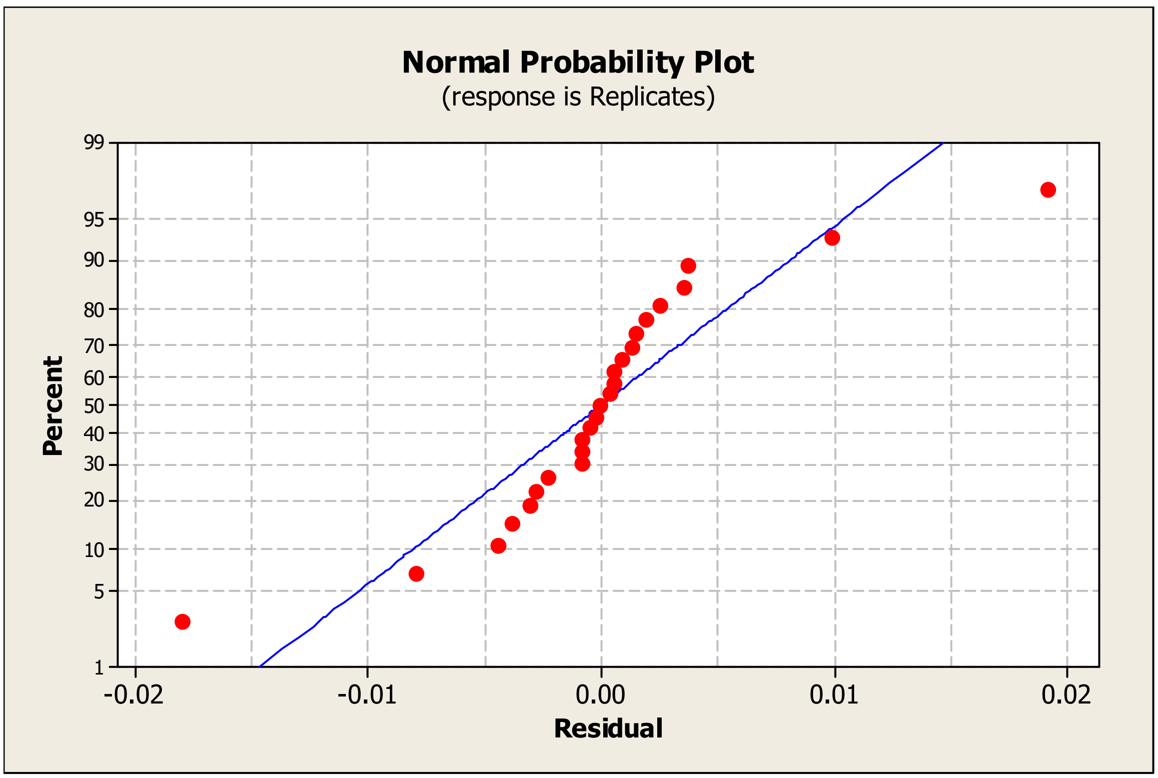
Since A, B, AB and curvature are significant (P<0.05), they have to be included in the regression model. There is also evidence of pure quadratic curvature.

Using Minitab:
Response Surface Regression: Replicates versus A, B

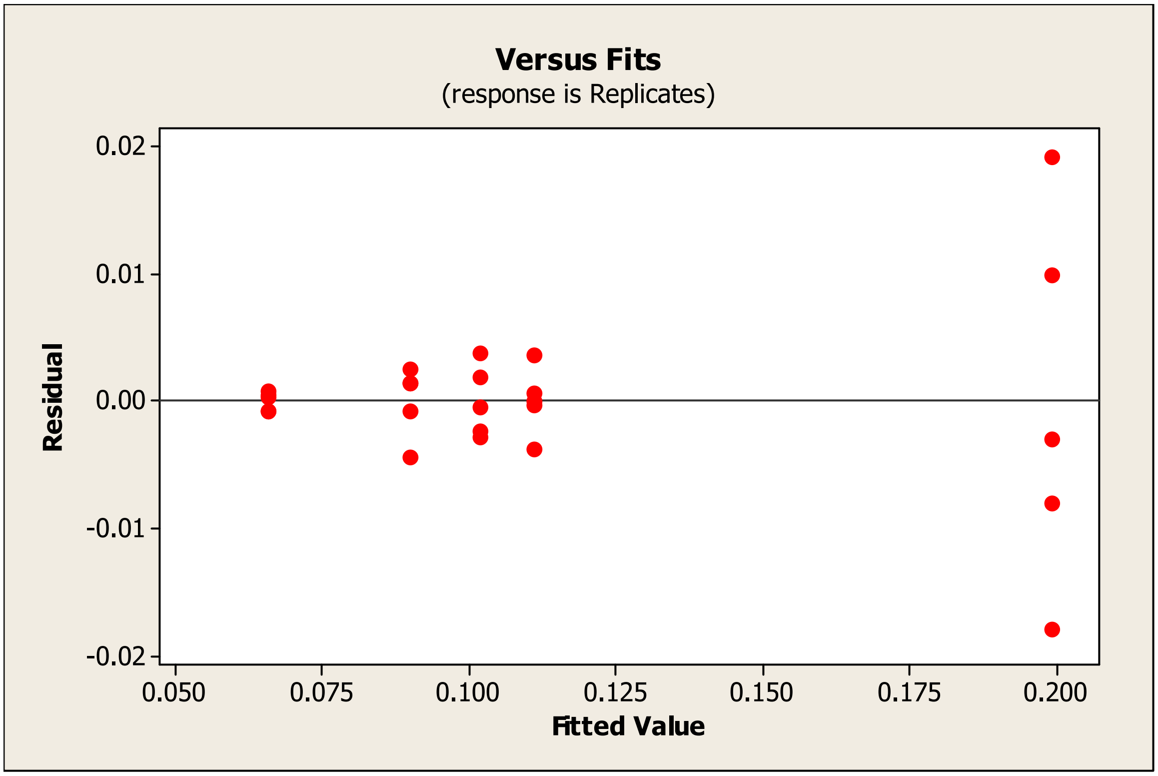
```
The following terms cannot be estimated, and were removed.

B*B


The analysis was done using coded units.

Estimated Regression Coefficients for Replicates

Term           Coef   SE Coef        T      P
Constant    0.10190  0.003083   33.053  0.000
A          -0.03862  0.001541  -25.054  0.000
B          -0.02813  0.001541  -18.249  0.000
A*A         0.01463  0.003447    4.244  0.000
A*B         0.01608  0.001541   10.432  0.000


S = 0.00689372  PRESS = 0.00148511
R-Sq = 98.19%   R-Sq(pred) = 97.18%  R-Sq(adj) = 97.83%


Analysis of Variance for Replicates

Source          DF    Seq SS    Adj SS    Adj MS       F      P
Regression       4  0.051684  0.051684  0.012921  271.88  0.000
  Linear         2  0.045656  0.045656  0.022828  480.35  0.000
  Square         1  0.000856  0.000856  0.000856   18.02  0.000
  Interaction    1  0.005171  0.005171  0.005171  108.82  0.000
Residual Error  20  0.000950  0.000950  0.000048
  Pure Error    20  0.000950  0.000950  0.000048
Total           24  0.052634
```

```
Unusual Observations for Replicates

Obs  StdOrder  Replicates    Fit  SE Fit  Residual  St Resid
  6         6       0.219  0.199   0.003     0.019      3.10 R
 16        16       0.181  0.199   0.003    -0.018     -2.91 R

R denotes an observation with a large standardized residual.


Estimated Regression Coefficients for Replicates using data in uncoded units

Term            Coef
Constant    0.101900
A         -0.0386200
B         -0.0281300
A*A        0.0146300
A*B        0.0160800
```

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{12} x_1 x_2$
$= 0.101900 - 0.0386200x_1 - 0.0281300x_2 + 0.0146300x_1^2 + 0.0160800\, x_1 x_2$



The normal probability plot looks skewed, with very little data lying on the blue line. The residuals do not seem to be following a normal distribution.



Variances of the residuals is shown to grow with increasing fitted values. However, the residuals are equally distributed above and below the center line.

b. Using transformed data sets exp(y):

From Minitab:
Response Surface Regression: Replicates versus A, B

```
The following terms cannot be estimated, and were removed.

B*B


The analysis was done using coded units.

Estimated Regression Coefficients for Replicates

Term           Coef   SE Coef        T      P
Constant    1.10728  0.003733  296.581  0.000
A          -0.04396  0.001867  -23.550  0.000
B          -0.03236  0.001867  -17.337  0.000
A*A         0.01779  0.004174    4.262  0.000
A*B         0.01933  0.001867   10.357  0.000


S = 0.00834830   PRESS = 0.00217794
R-Sq = 98.00%    R-Sq(pred) = 96.88%   R-Sq(adj) = 97.60%


Analysis of Variance for Replicates

Source          DF    Seq SS    Adj SS    Adj MS       F      P
Regression       4  0.068342  0.068342  0.017086  245.15  0.000
  Linear         2  0.059599  0.059599  0.029800  427.58  0.000
```

Square	1	0.001266	0.001266	0.001266	18.17	0.000
Interaction	1	0.007477	0.007477	0.007477	107.28	0.000
Residual Error	20	0.001394	0.001394	0.000070		
Pure Error	20	0.001394	0.001394	0.000070		
Total	24	0.069736				

Unusual Observations for Replicates

Obs	StdOrder	Replicates	Fit	SE Fit	Residual	St Resid
6	6	1.244	1.221	0.004	0.023	3.14 R
16	16	1.199	1.221	0.004	-0.022	-2.92 R

R denotes an observation with a large standardized residual.

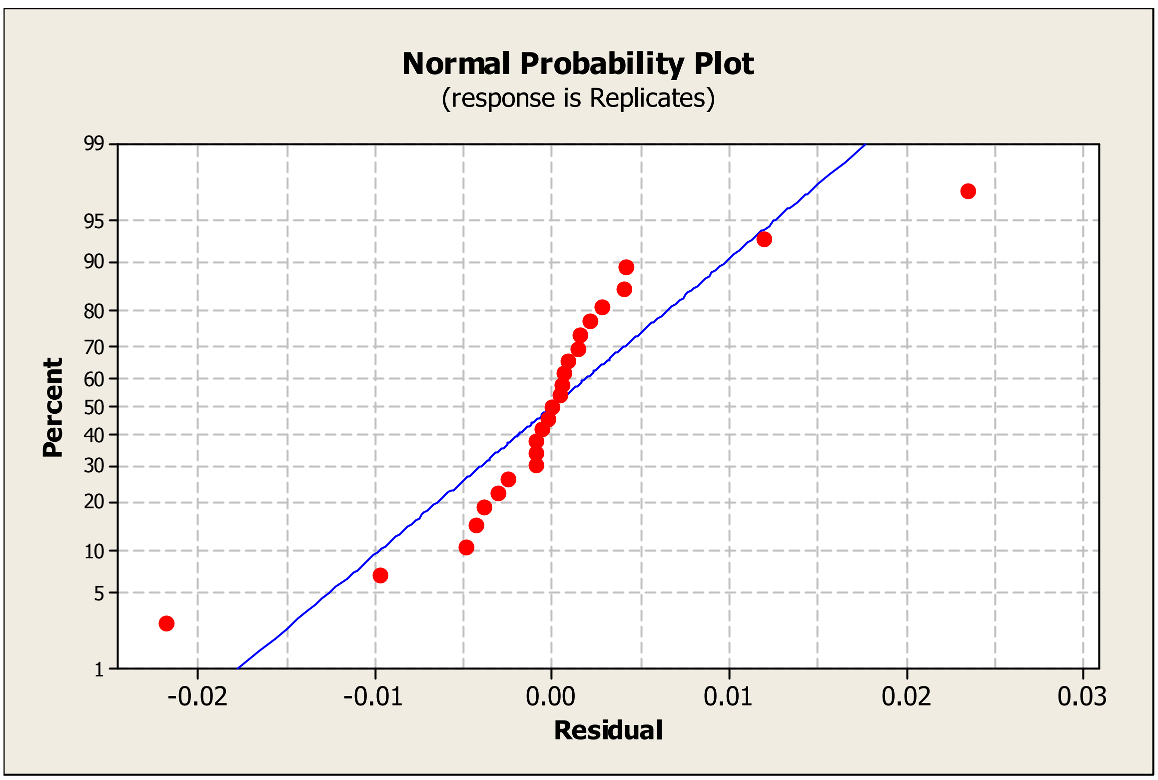
Estimated Regression Coefficients for Replicates using data in uncoded units

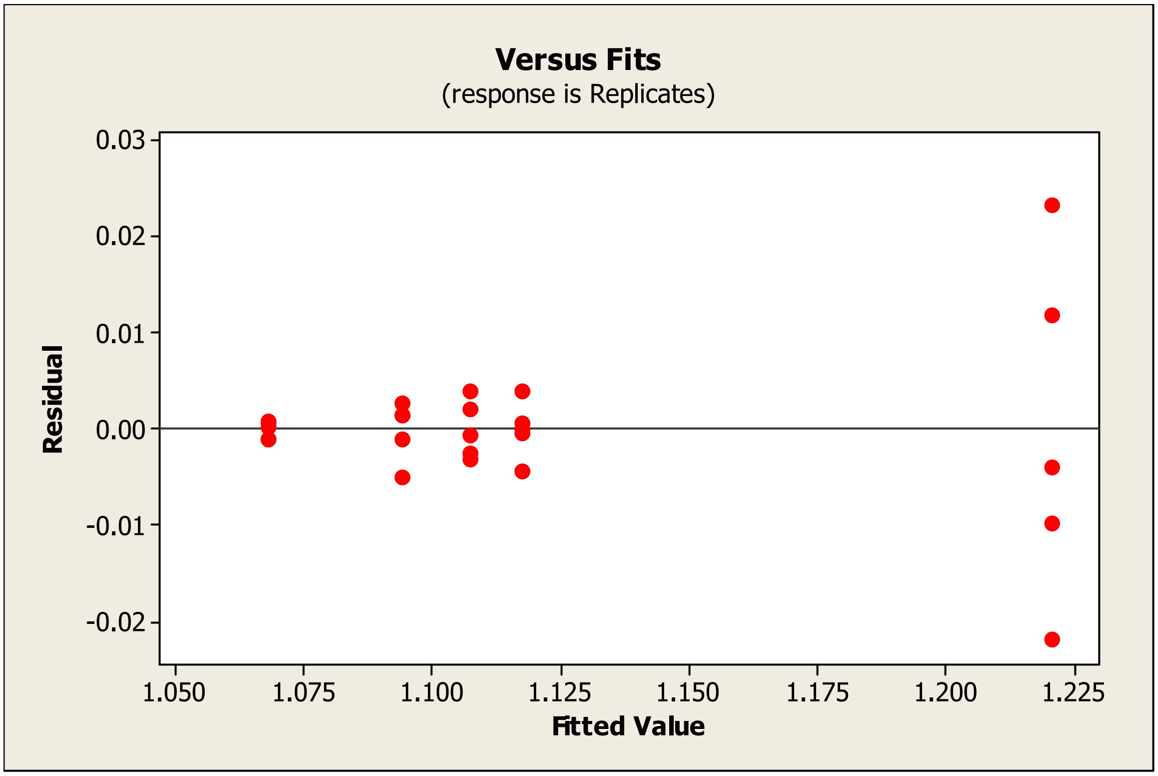
Term	Coef
Constant	1.10728
A	-0.0439614
B	-0.0323629
A*A	0.0177910
A*B	0.0193347

Regression model:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{11} x_1^2 + \beta_{12} x_1 x_2$

$= 1.10728 - 0.0439614x_1 - 0.0323629x_2 + 0.0177910x_1^2 + 0.0193347x_1x_2$





Both plots do not show much improvement from the previous plots.

Using transformation 1/y:
From Minitab:
Response Surface Regression: Replicates versus A, B

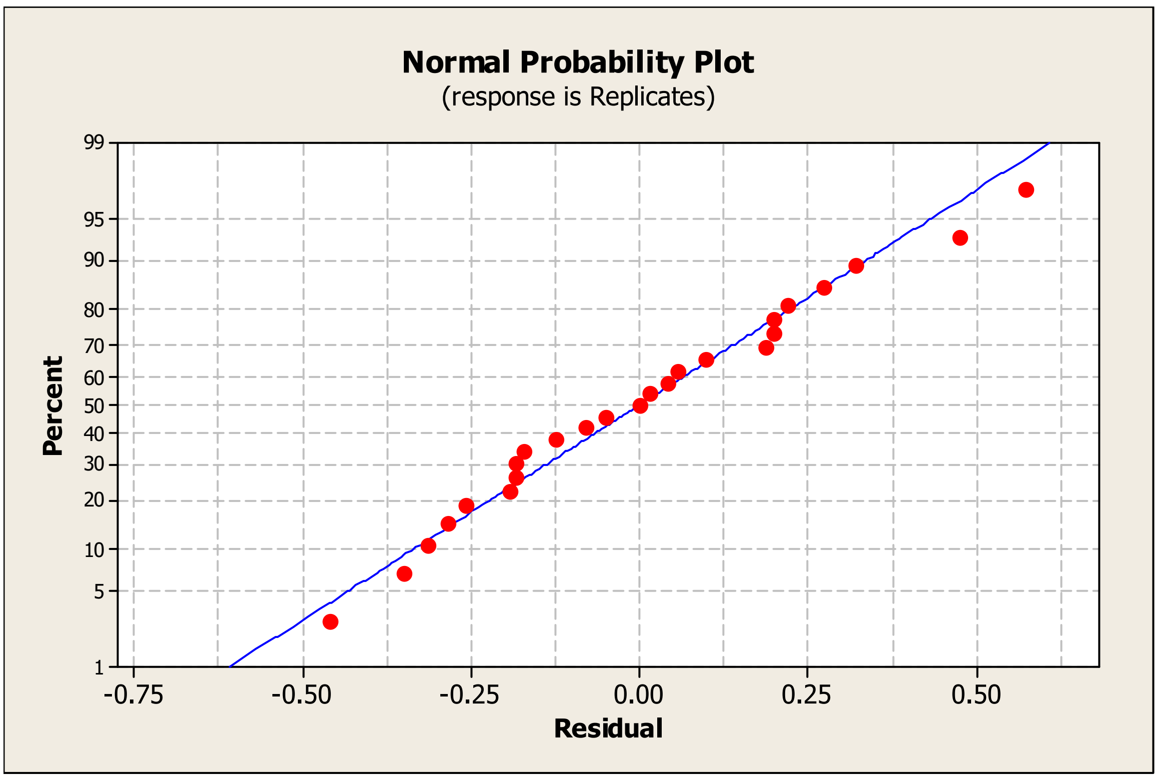
```
The following terms cannot be estimated, and were removed.

B*B


The analysis was done using coded units.

Estimated Regression Coefficients for Replicates

Term          Coef  SE Coef       T      P
Constant   9.81930  0.12827  76.551  0.000
A          3.06367  0.06414  47.769  0.000
B          2.01027  0.06414  31.344  0.000
A*A        0.27219  0.14341   1.898  0.072
A*B        0.02012  0.06414   0.314  0.757


S = 0.286823   PRESS = 2.57085
R-Sq = 99.39%  R-Sq(pred) = 99.05%  R-Sq(adj) = 99.27%


Analysis of Variance for Replicates

Source          DF   Seq SS   Adj SS   Adj MS        F      P
Regression       4  268.849  268.849   67.212   817.00  0.000
  Linear         2  268.545  268.545  134.272  1632.15  0.000
  Square         1    0.296    0.296    0.296     3.60  0.072
  Interaction    1    0.008    0.008    0.008     0.10  0.757
Residual Error  20    1.645    1.645    0.082
```

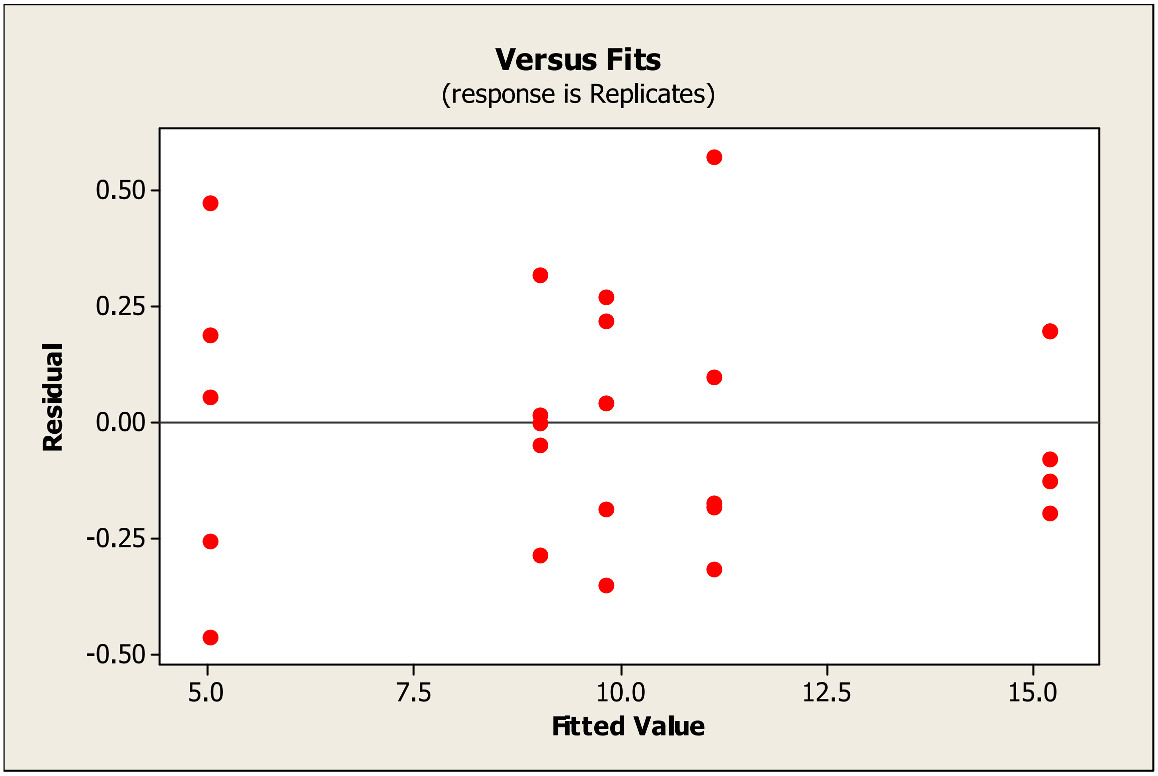
```
  Pure Error    20    1.645    1.645    0.082
Total           24  270.495


Unusual Observations for Replicates

Obs  StdOrder  Replicates      Fit  SE Fit  Residual  St Resid
 17        17      11.696   11.125   0.128     0.571      2.23 R

R denotes an observation with a large standardized residual.


Estimated Regression Coefficients for Replicates using data in uncoded units

Term            Coef
Constant     9.81930
A            3.06367
B            2.01027
A*A         0.272187
A*B        0.0201165
```

Regression model:

$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{11} x_1^2 + \beta_{12} x_1 x_2$

$= 9.81930 + 3.063674x_1 + 2.01027x_2 + 0.272187x_1^2 + 0.0201165 x_1 x_2$





Comparing these two new plots, there is much improvements in the sense that the residual VS fitted value plots do not show a growth in variance. Also, the normal probability plot shows a more well fitted data to line, hence randomly distributed data. Thus, the last transformation 1/y seems more appropriate for fitting the current regression model.

Problem 4 (13-12)

13-12.

Response Surface Regression: y versus x1, x2, z

```
The analysis was done using coded units.

Estimated Regression Coefficients for y

Term         Coef  SE Coef      T     P

Constant   87.3333    1.681  51.968  0.000

x1         9.8013    1.873   5.232  0.001

x2         2.2894    1.873   1.222  0.256

z         -6.1250    1.455  -4.209  0.003

x1*x1    -13.8333    3.361  -4.116  0.003

x2*x2    -21.8333    3.361  -6.496  0.000

z*z        0.1517    2.116   0.072  0.945

x1*x2      8.1317    4.116   1.975  0.084

x1*z      -4.4147    2.448  -1.804  0.109

x2*z      -7.7783    2.448  -3.178  0.013

…

Analysis of Variance for y

Source         DF   Seq SS   Adj SS   Adj MS     F     P

Regression      9  2034.94  2034.94  226.105  13.34  0.001

  Linear        3   789.28   789.28  263.092  15.53  0.001

  Square        3   953.29   953.29  317.764  18.75  0.001

  Interaction   3   292.38   292.38   97.458   5.75  0.021

Residual Error  8   135.56   135.56   16.945

  Lack-of-Fit   3    90.22    90.22   30.074   3.32  0.115

  Pure Error    5    45.33    45.33    9.067

Total          17  2170.50

…

Estimated Regression Coefficients for y using data in uncoded units

Term        Coef

Constant  87.3333

x1        5.8279

x2        1.3613

z        -6.1250

x1*x1    -4.8908
```

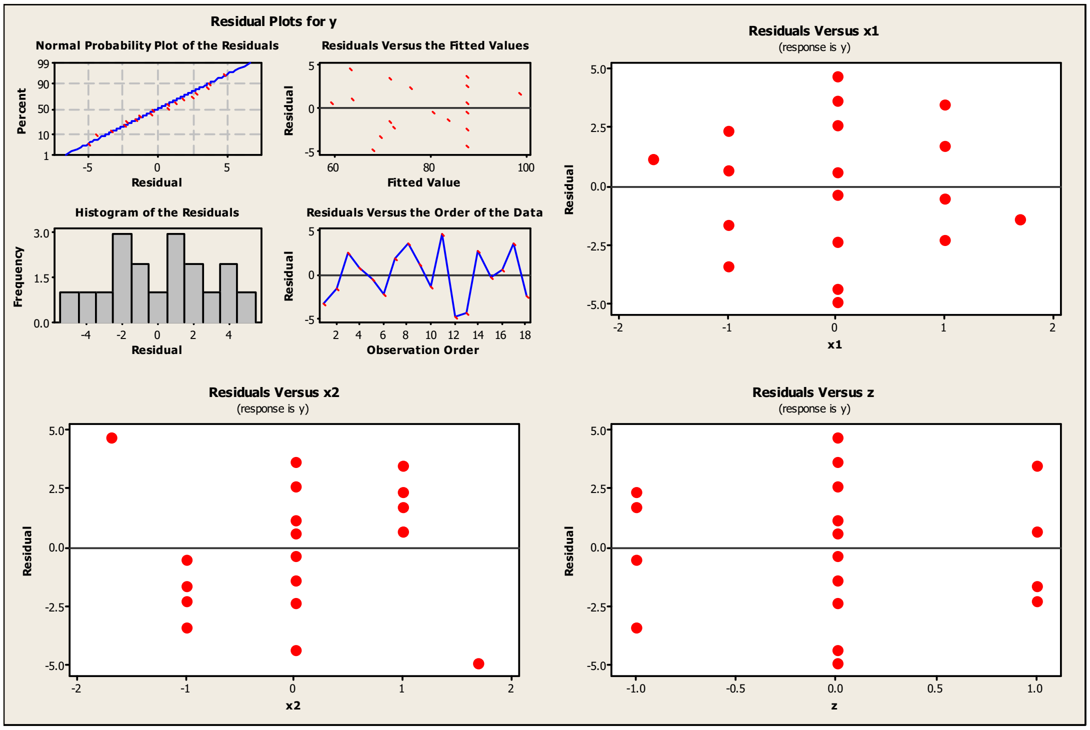
```
x2*x2    -7.7192
z*z       0.1517
x1*x2     2.8750
x1*z     -2.6250
x2*z     -4.6250
```

The coefficients for x_1z and x_2z (the two interactions involving the noise variable) are significant (P-values ≤ 0.10), so there is a robust design problem.

Reduced model:

Response Surface Regression: y versus x1, x2, z

```
The analysis was done using coded units.
Estimated Regression Coefficients for y
Term        Coef  SE Coef      T     P
Constant   87.361    1.541  56.675  0.000
x1          9.801    1.767   5.548  0.000
x2          2.289    1.767   1.296  0.227
z          -6.125    1.373  -4.462  0.002
x1*x1     -13.760    3.019  -4.558  0.001
x2*x2     -21.760    3.019  -7.208  0.000
x1*x2       8.132    3.882   2.095  0.066
x1*z       -4.415    2.308  -1.912  0.088
x2*z       -7.778    2.308  -3.370  0.008
…
Analysis of Variance for y
Source          DF   Seq SS   Adj SS   Adj MS     F     P
Regression       8  2034.86  2034.86  254.357  16.88  0.000
  Linear         3   789.28   789.28  263.092  17.46  0.000
  Square         2   953.20   953.20  476.602  31.62  0.000
  Interaction    3   292.38   292.38   97.458   6.47  0.013
Residual Error   9   135.64   135.64   15.072
  Lack-of-Fit    4    90.31    90.31   22.578   2.49  0.172
  Pure Error     5    45.33    45.33    9.067
Total           17  2170.50
…
```



$y_{\text{Pred}} = 87.36 + 5.83x_1 + 1.36x_2 - 4.86x_1^{\ 2} - 7.69x_2^{\ 2} + (-6.13 - 2.63x_1 - 4.63x_2)z$

For the mean yield model, set $z = 0$:

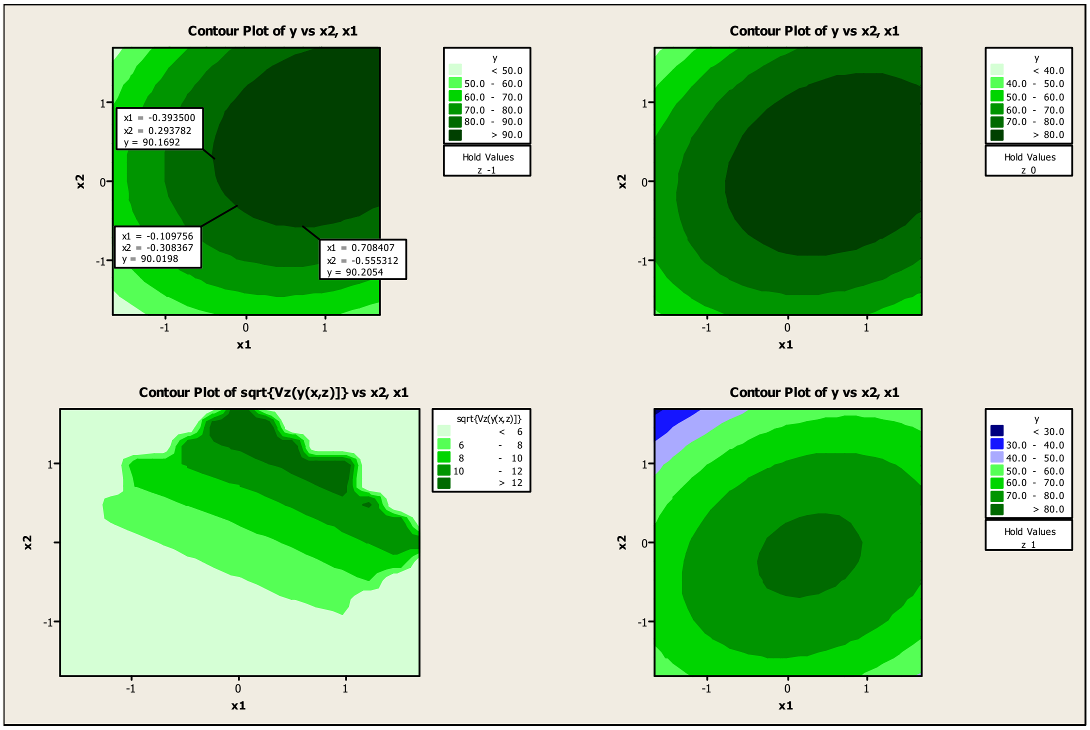
Mean Yield = $87.36 + 5.83x_1 + 1.36x_2 - 4.86x_1^{\ 2} - 7.69x_2^{\ 2}$

For the variance model, assume $\sigma_z^{\ 2} = 1$:

Variance of Yield = $\sigma_z^{\ 2}\,(-6.13 - 2.63x_1 - 4.63x_2)^2 + \hat{\sigma}^2$

$= (-6.13 - 2.63x_1 - 4.63x_2)^2 + 15.072$

This equation can be added to the worksheet and used in a contour plot with x_1 and x_2.



Examination of contour plots for Free Height show that heights greater than 90 are achieved with $z = -1$. Comparison with the contour plot for variability shows that growth greater than 90 with minimum variability is achieved at approximately $x_1 = -0.11$ and $x_2 = -0.31$ (mean yield of about 90 with a standard deviation between 6 and 8). There are other combinations that would work.

Note: the question was unclear as to whether the noise input z was controllable. If so, selecting z = -1 may give minimal sensitivity of the output to variation in z. If, however, we assume that z cannot be controlled, we must assume it to have zero mean and constant variance. The alternative solution following (courtesy H. Hu) shows a solution based on the assumption that z cannot be controlled.

Problem 4

Montgomery 13-12

Reconsider the crystal growth experiment from Exercise 13-10. Suppose that $x_3 = z$ is now a noise variable, and that the modified experimental design shown here has been conducted. The experimenters want the growth rate to be as large as possible but they also want the variability transmitted from z to be small. Under what set of conditions is growth greater than 90 with minimum variability achieved?

x_1	x_2	z	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	0	113

0	0	0	100
0	0	0	118
0	0	0	88
0	0	0	100
0	0	0	85

We use Minitab to do the robustness study. The experimental design is a “modified” central composite design in which the axial runs in the z direction have been eliminated.

```
Response Surface Regression: response versus x1, x2, x3

The analysis was done using coded units.

Estimated Regression Coefficients for response

Term         Coef  SE Coef       T      P
Constant   98.896    5.607  17.639  0.000
x1          1.271    3.821   0.333  0.747
x2          1.361    3.821   0.356  0.730
x3         -6.125    4.992  -1.227  0.251
x1*x1      -5.412    3.882  -1.394  0.197
x2*x2     -14.074    3.882  -3.625  0.006
x1*x2       2.875    4.992   0.576  0.579
x1*x3      -2.625    4.992  -0.526  0.612
x2*x3      -4.625    4.992  -0.926  0.378


S = 14.1209    PRESS = 9196.84
R-Sq = 65.33%  R-Sq(pred) = 0.00%  R-Sq(adj) = 34.51%


Analysis of Variance for response

Source          DF  Seq SS  Adj SS   Adj MS     F      P
Regression       8  3381.2  3381.2   422.65  2.12  0.142
  Linear         3   347.5   347.5   115.84  0.58  0.642
  Square         2  2741.3  2741.3  1370.65  6.87  0.015
  Interaction    3   292.4   292.4    97.46  0.49  0.699
```

Residual Error	9	1794.6	1794.6	199.40		
Lack-of-Fit	4	935.3	935.3	233.81	1.36	0.365
Pure Error	5	859.3	859.3	171.87		
Total	17	5175.8				

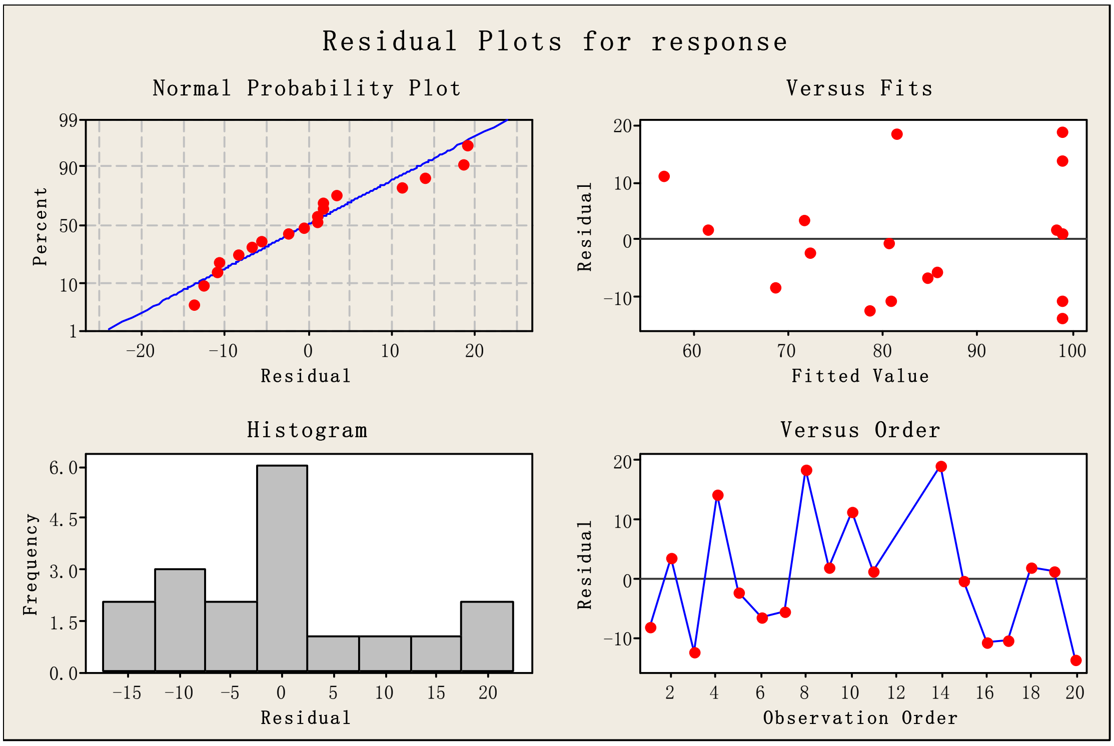
Unusual Observations for response

Obs	StdOrder	response	Fit	SE Fit	Residual	St Resid
8	9	100.000	81.449	10.836	18.551	2.05 R

R denotes an observation with a large standardized residual.

Estimated Regression Coefficients for response using data in uncoded units

Term	Coef
Constant	98.8959
x1	1.27146
x2	1.36130
x3	-6.12500
x1*x1	-5.41231
x2*x2	-14.0744
x1*x2	2.87500
x1*x3	-2.62500
x2*x3	-4.62500



The response model for the process robustness study is :

$$y(x,z)=f(x)+h(x,z)+\varepsilon$$

$$=\beta_0+\beta_1 x_1+\beta_2 x_2+\beta_{11}x_1^2+\beta_{22}x_2^2+\beta_{12}x_1x_2+\gamma_1 z+\delta_{11}x_1 z+\delta_{21}x_2 z+\varepsilon$$

$$\hat{y}(x,z)=98.8959+1.27146x_1+1.36130x_2-5.41231x_1^2-14.0744x_2^2+2.875x_1x_2-6.125z-2.625x_1z-4.625x_2z$$

Therefore the mean model is

$$E_z[y(x,z)]=f(x)=98.8959+1.27146x_1+1.36130x_2-5.41231x_1^2-14.0744x_2^2+2.875x_1x_2$$

The variance model is

$$Vz[y(x,z)]=\sigma_z^2\left(\frac{\partial h(x,z)}{\partial z}\right)^2+\sigma^2$$

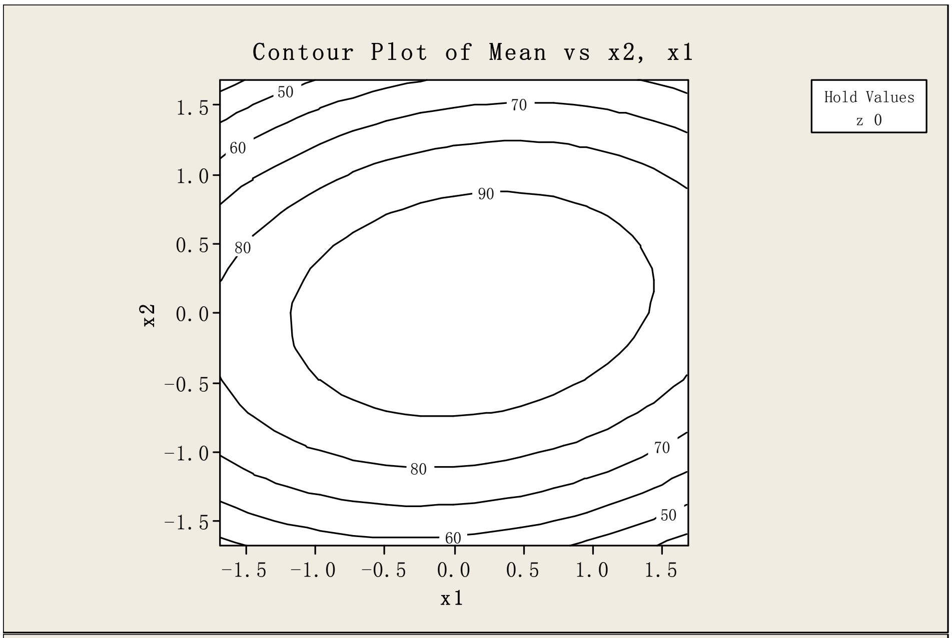
$$=\sigma_z^2(-6.125-2.625x_1-4.625x_2)^2+\sigma^2$$

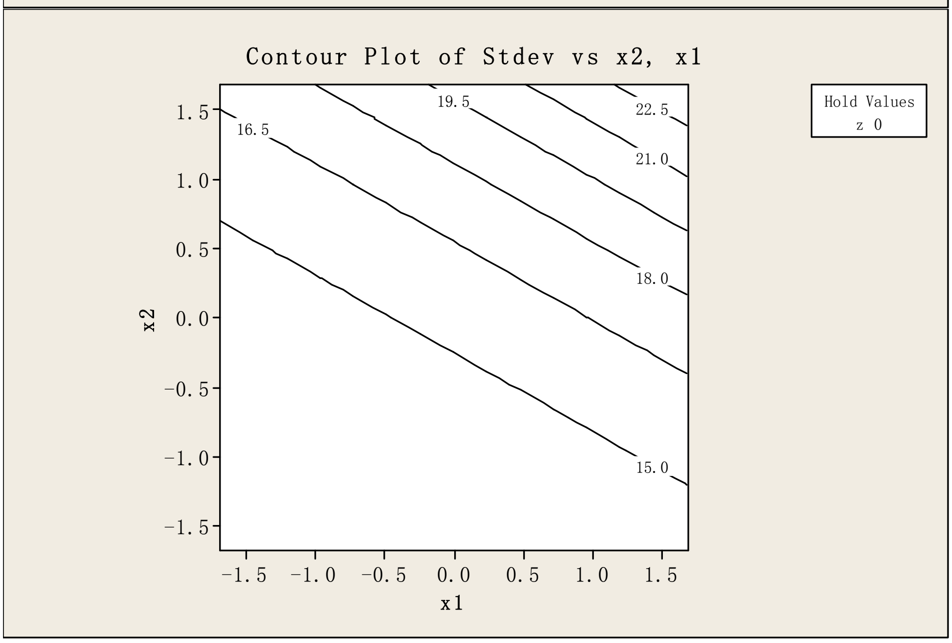
Now we assume that the low and high levels of the noise variable z have been run at one standard deviation either side of its typical or average value, so that $\sigma_z^2=1$ and since the residual mean square from fitting the response model is $MS_E=199.40$will use $\hat{\sigma}^2=MS_E=199.40$

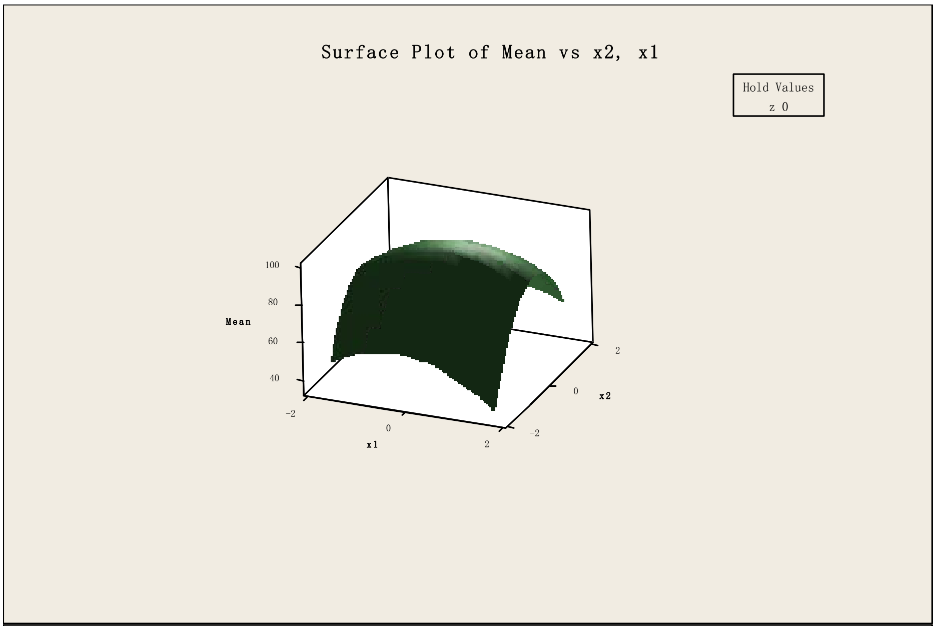
Therefore the variance model

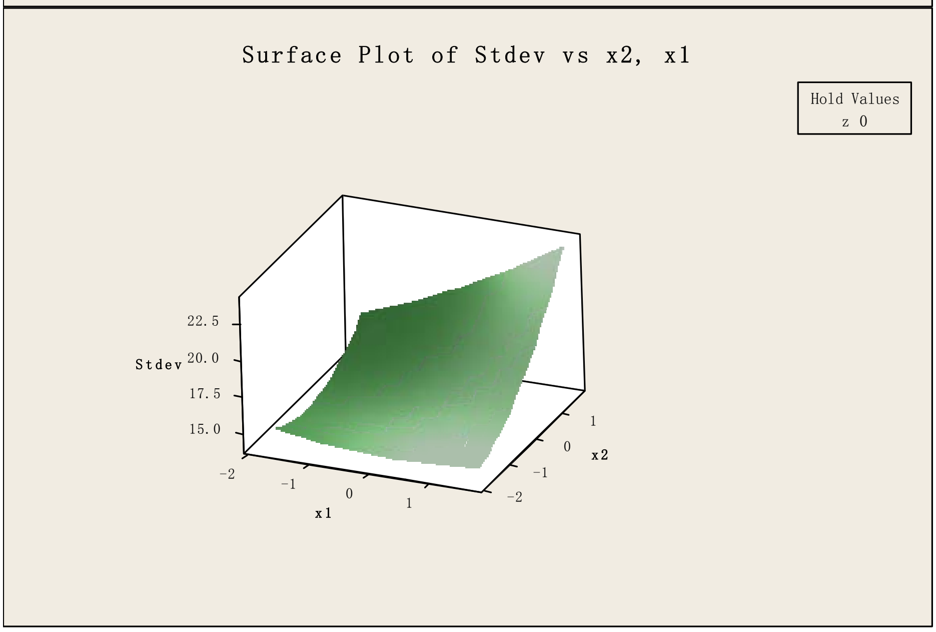
$$Vz[y(x,z)]= (-6.125-2.625x_1-4.625x_2)^2+199.40$$

Following figures show response surface contour plots and three-dimensional surface plots of the mean model and the standard deviation respectively.

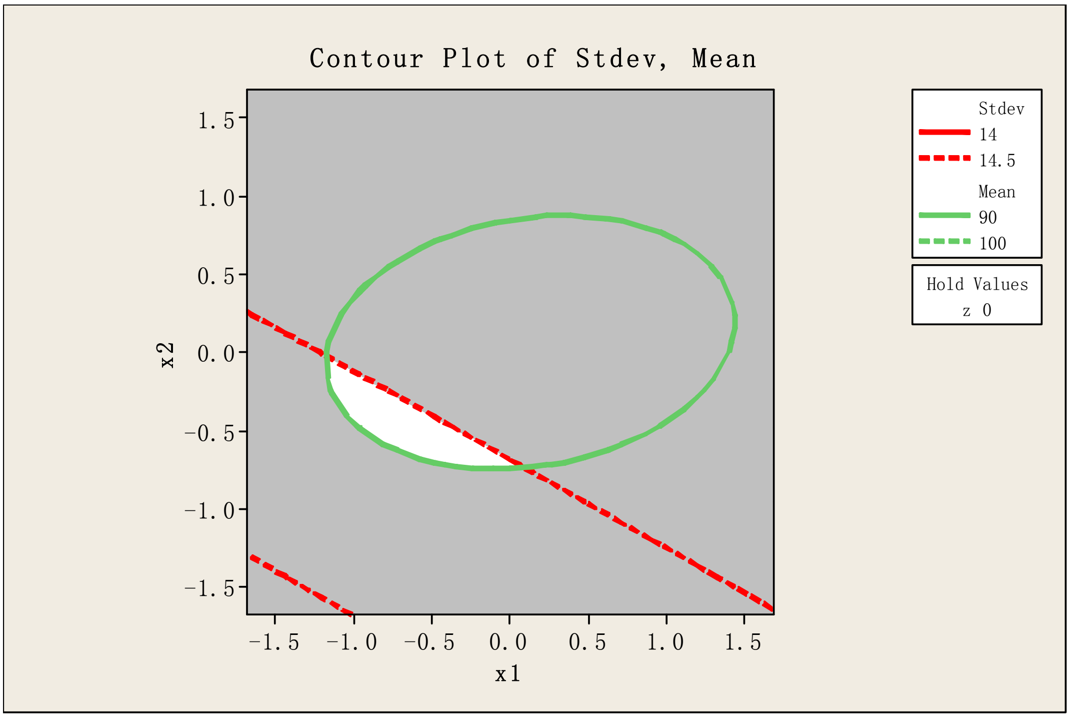








The objective of the robustness study is to find a set of operating conditions that would result in a mean response greater than 90 from the mean model with the minimum contour of standard deviation. The unshaded region of the following plot indicates operating conditions on x_1 and x_2, where the requirements for the mean response larger than 90 are satisfied and the response standard deviation do not exceed 14.5.



Actually, if we use Excel Solver, we can get a optimal solution for minimizing the standard deviation with the constraint that the mean value is greater than 90.

The optimal solution is : mean=90

Stdv=14.18

X_1=-0.83 x_2=-0.58

This solution conforms to the analysis we did above.