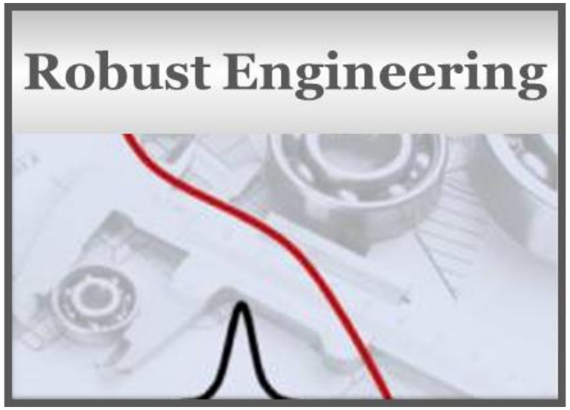
Statistiques Appliquées Formation Industrielle

Module: Robust Engineering

Design of Experiments & Response Surface Modelling



Session 6: Analysing a Two-Level Screening Experiment

Reminder of Our Earlier Analysis Sessions

- In Sessions 2 and 3, we analysed data from a sintering process
  - we used multiple regression to generate a response surface.
  - we plotted the response surface.
  - we constructed an ANOVA table and calculated R^2, also PRESS RMSE.
  - we looked at plots of residuals.
  - we used t tests on individual regression coefficients to identify factors that have a real effect on the response.

In This Session We Will…

- Analyse the data from a small screening experiment based on a full factorial 2^3 design.
- Show that the special features of this design make the multiple regression calculations much simpler.
- Define and interpret two new types of plot:
  - interaction plots and main effect plots.
- Explain that although statistical software will produce the usual regression output, R^2, PRESS and residual plots are of little or no use when the residual degrees of freedom (ResDF) are small.
- Explain that t tests will have low power unless the design is replicated.

Back To Our Bumper Bracket Study

- One of the functions of the part is to support the mass of the bumper below the grill opening, and the under-tray mounted on the lower spoiler.
- It also has to;
  - resist aerodynamic loading on bumper and under-tray surfaces.
  - give predictable strength during pedestrian impact.
  - prevent significant damage during 'snow bank' impact test.
- All these functions require the material to have predictable strength.
- The material under study was glass-reinforced polypropylene.

Back To Our Bumper Bracket Study (cont.)

- One objective of the study was to assess whether the use of re-ground material would affect tensile strength (due to reduction in the length of the glass fibres and the molecular weight of the polymer) under a range of processing conditions

A 2^3 Experiment

Factor	**Name**	**Units**	**–1 level**	**+1 level**
x_1	Melt temp	℃	230	270
x_2	Screw speed	rpm	50	300
x_3	Hold pressure	MPa	50	250

In this experiment, 10% of the material was re-cycled

x_1	x_2	x_3	Tensile strength (MPa)
–1	–1	–1	40.49
+1	–1	–1	39.94
–1	+1	–1	40.07
+1	+1	–1	39.86
–1	–1	+1	42.78
+1	–1	+1	41.74
–1	+1	+1	42.56
+1	+1	+1	42.49

Sample mean from 3 parts

Coded design and response values

Generating The Response Surface

- We will fit a response surface equation including all the 2fi's:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3$$

- The regression coefficients can be found from the usual formula:

$$\underline{b} = (M'M)^{-1}M'\underline{y}$$

  - where the model matrix M has a column for every term we want to fit.

we will look at the model matrix on the next slide

- In coded factor levels, the fitted equation is:

$$y = 41.24 - 0.23\ x_1 + 0.00\ x_2 + 1.15x_3 + 0.16\ x_1x_2 - 0.04\ x_1x_3 + 0.13\ x_2x_3$$

Not actually 0, but very close

In an orthogonal two-level experiment like this one, there is a much simpler way to calculate the coefficients

Calculating The Regression Coefficients From Effects

- For each column in M apart from the intercept:
  - pick out the rows with a '+1' and average the y values;
  - pick out the rows with a '−1' and average the y values;
  - subtract the '−1' average from the '+1' average;

This step gives us the effect for each column

  - divide the effect by 2.

x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	y
-1	-1	-1	+1	+1	+1	40.49
+1	–1	–1	–1	–1	+1	39.94
–1	+1	–1	–1	+1	–1	40.07
+1	+1	–1	+1	–1	–1	39.86
–1	–1	+1	+1	–1	–1	42.78
+1	–1	+1	–1	+1	–1	41.74
–1	+1	+1	–1	–1	+1	42.56
+1	+1	+1	+1	+1	+1	42.49

For the x1 column we get:

Ave (+)	41.01					
Ave (–)	41.48					
Effect	-0.47					
Coefficient	-0.23					

Calculating The Regression Coefficients From Effects

- For each column in M apart from the intercept:
  - pick out the rows with a '+1' and average the y values;
  - pick out the rows with a '−1' and average the y values;
  - subtract the '−1' average from the '+1' average;
  - divide the effect by 2.

x_1	x_2	x_3	x_1x_2	x_1x_3	x_2x_3	y
-1	-1	-1	+1	+1	+1	40.49
+1	–1	–1	–1	–1	+1	39.94
–1	+1	–1	–1	+1	–1	40.07
+1	+1	–1	+1	–1	–1	39.86
–1	–1	+1	+1	–1	–1	42.78
+1	–1	+1	–1	+1	–1	41.74
–1	+1	+1	–1	–1	+1	42.56
+1	+1	+1	+1	+1	+1	42.49

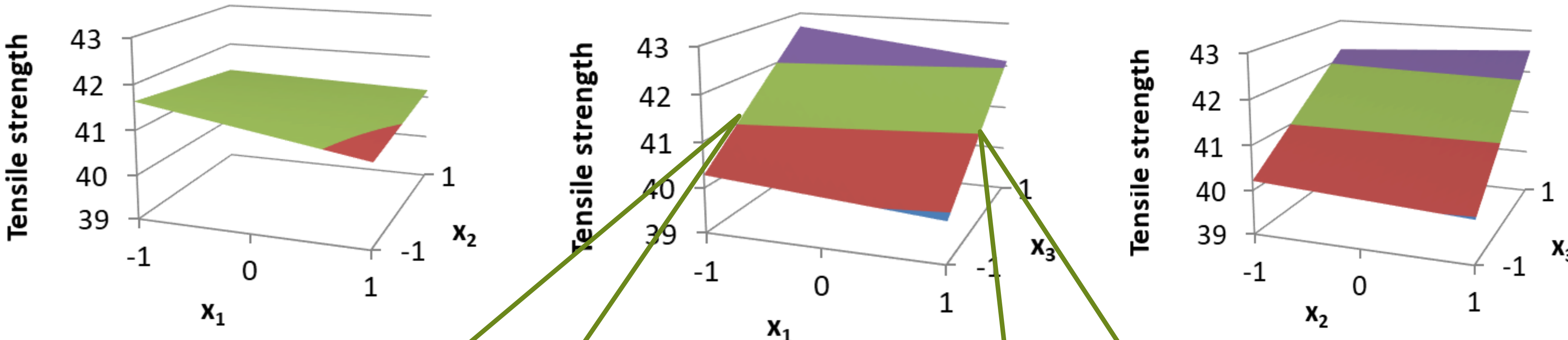
Complete set of calculations

Ave (+)	41.01	41.25	42.39	41.41	41.20	41.37
Ave (–)	41.48	41.24	40.09	41.08	41.29	41.11
Effect	-0.47	0.01	2.30	0.33	-0.09	0.26
Coefficient	-0.23	0.00	1.15	0.16	-0.04	0.13

Plotting The Response Surface

- We can make multiple plots of the surface, as before
  - and the plots show why we divide an effect by 2

In each plot the 3rd factor is set at coded 0



Average strength when $x_1 = -1$ is 41.48

Average strength when $x_1 = +1$ is 41.01

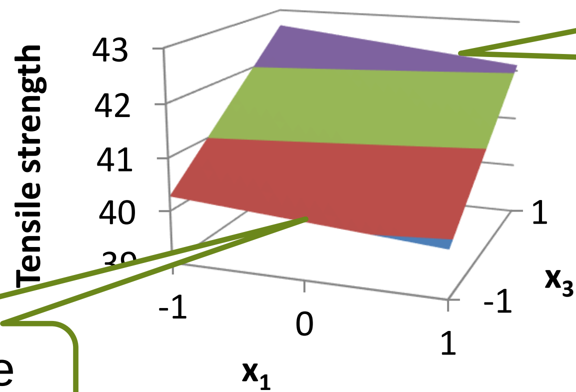
Difference = -0.47 (the effect of x_1)
Gradient across the middle of the surface is $= -0.47/2 = -0.23 = b_1$

The linear coefficient for x_1

So b_1 can be interpreted as the gradient across the middle of the surface

Interpreting A 2fi Coefficient

- Using the same kind of geometrical argument, we can show that a 2fi coefficient is half the difference between the gradients at the two edges of the surface:



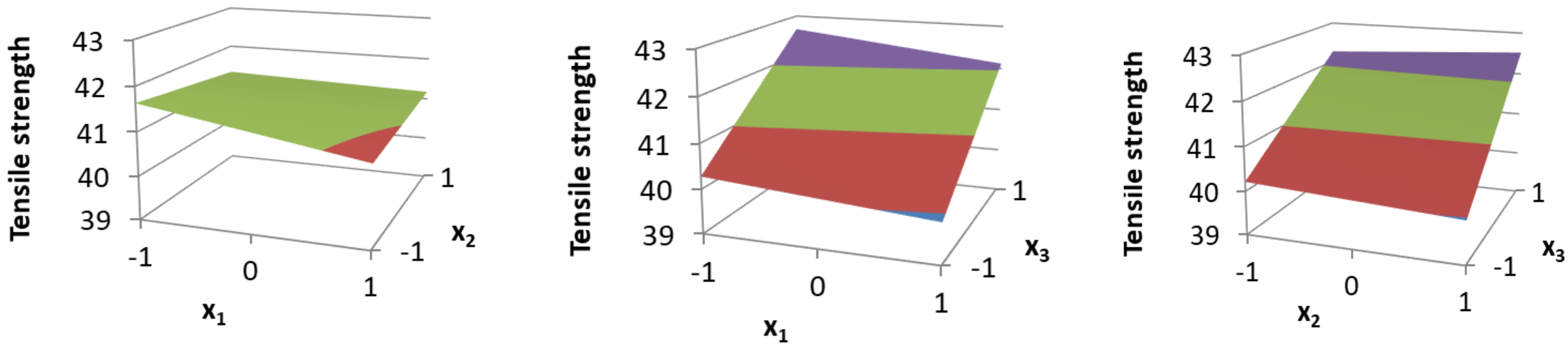
Gradient at the back is $-\ 0.2775$

Gradient at the front is $-\ 0.1900$

Half the difference is
$(-\ 0.2775\ +\ 0.1900) = -0.04375$

…. and this was the value of b_{13}, the 2fi coefficient for x_1x_3

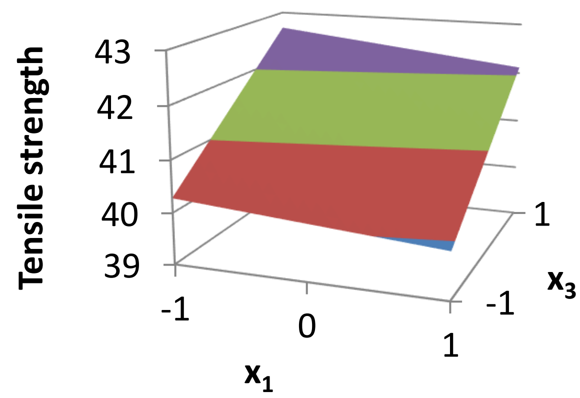
Plotting The Response Surface (cont.)

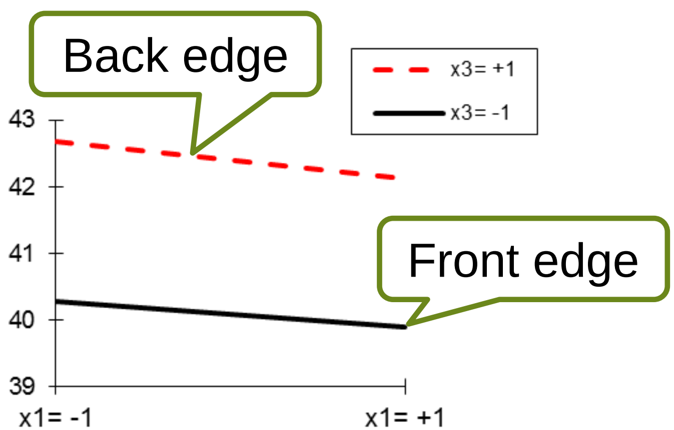


- The number of possible 3D plots becomes unmanageable for more than 4 or 5 factors
  - but there is a simpler way of looking at the surface, using interaction plots.

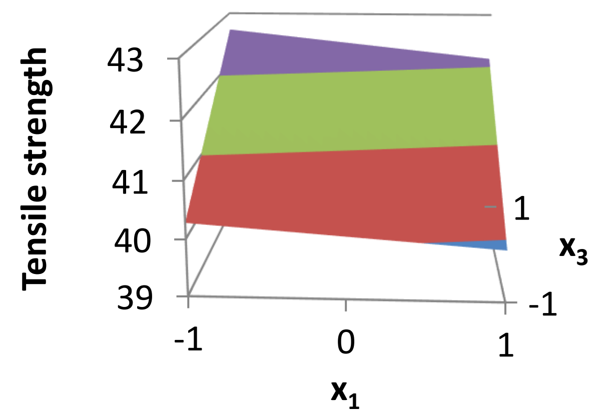
Example of An Interaction Plot

- The plot on the right-bottom shows the surface 'side on', looking at right-angles to the x_1 axis.
- An interaction plot shows just the front and back edges of the surface, plotted against a single scale.





Parallel lines imply no twist i.e. no interaction



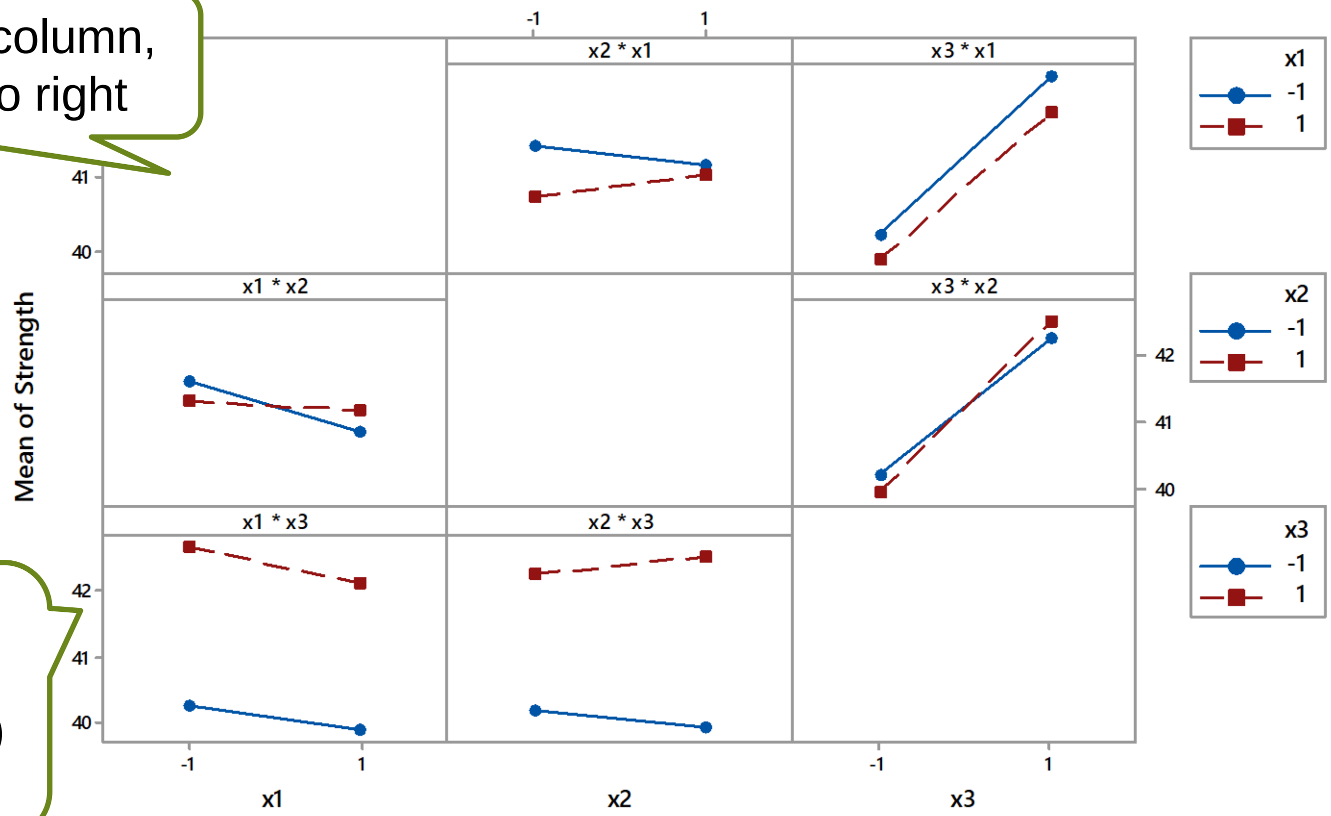
Interaction Plots

- Minitab can display two interaction plots for each pair of factors.

For all plots in this column, x_1 goes from left to right

Top right & bottom left are two views of the same 3D surface

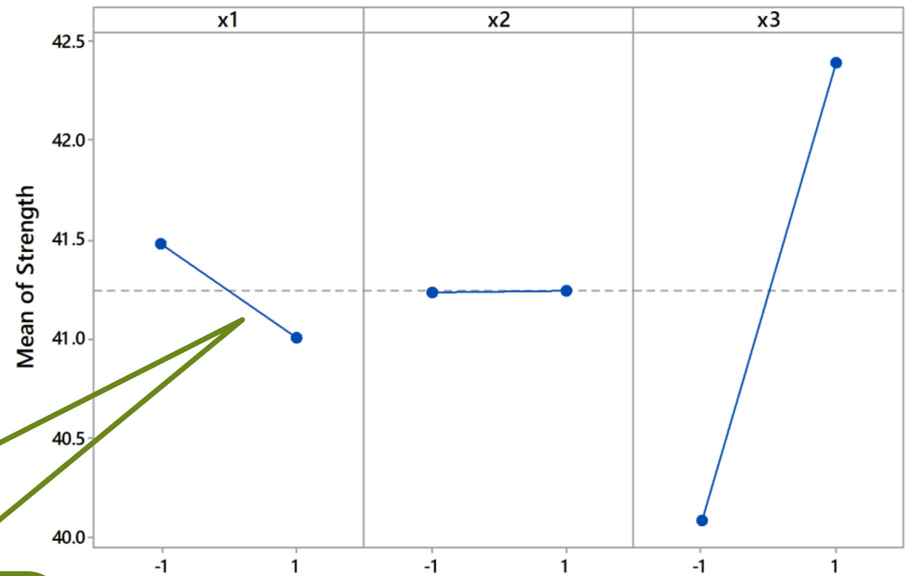
In this row the blue (solid) line is for $x_3 = -1$, red (dashed) line is for $x_3 = +1$



Main Effect Plot

- There is one more type of plot which can be constructed very easily in orthogonal two-level experiments.
  - for each factor, it shows the average response at -1 and at $+1$, as in our table earlier.

x_1	Melt temp
x_2	Screw speed
x_3	Holding pressure



Main effect plots must be interpreted cautiously

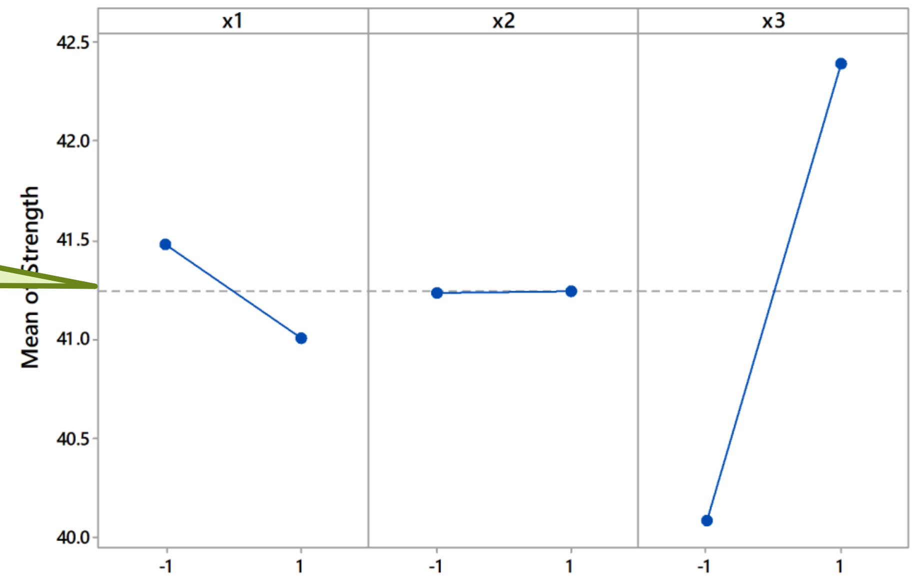
These are average gradients across the surface – only useful if interactions are small

	x_1	x_2	x_3
Ave (+)	41.01	41.25	42.39
Ave (–)	41.48	41.24	40.09

Main Effect Plot (cont.)

The horizontal line is at the overall mean response 41.24

This is our intercept in the regression equation



- Coded and uncoded factor levels:
  - we have shown the plots with coded scales because we think this helps you to understand how they work.
  - in a real application you will probably want to use uncoded scales.

ANOVA Table and R^2

Source of variation	Sum of squares	Degrees of freedom
Regression (RegSS)	11.403	6
Residual (ResSS)	0.050	1
Total (SSy)	11.452	7

terms less the intercept

n – # terms

$n - 1$

$$R^2 = \frac{RegSS}{SSy} = \frac{11.403}{11.452} = 0.996$$

This seems impressive!

…. but look at ResDF

When ResDF is very small, most data sets will have R^2 close to 1

What about residual plots?

- Here are the residuals for the 2^3 experiment:

Actual response	Fitted response	Residual
40.49	40.57	-0.08
39.94	39.86	0.08
40.07	39.99	0.08
39.86	39.94	-0.08
42.78	42.70	0.08
41.74	41.82	-0.08
42.56	42.64	-0.08
42.49	42.41	0.08

All the residuals are ±0.08!

When ResDF = 1, all the residuals will have the same absolute value

So we won't bother to plot the residuals!

PRESS

- PRESS is the sum of squares of the leave-one-out prediction errors.
- The formula for a leave-one-out prediction error is:

$$\frac{\text{residual}}{(1 - \text{leverage})}$$

- The leave-one-out prediction error is ± 0.63 for all 8 points.

These prediction errors are meaningless when ResDF is very small

Residual	Leverage	Prediction error
-0.08	0.875	-0.63
0.08	0.875	0.63
0.08	0.875	0.63
-0.08	0.875	-0.63
0.08	0.875	0.63
-0.08	0.875	-0.63
-0.08	0.875	-0.63
0.08	0.875	0.63

Selecting The Important Effects

- In session 3 we used t tests on the regression coefficients
- The results for our 2^3 example:

$$T = \frac{\text{b}}{\text{SE of b}}$$

Term	Effect	Coeff.	SE	T	P
x_1	-0.4675	-0.2338	0.07875	-2.97	0.207
x_2	0.0075	0.0038	0.07875	0.05	0.970
x_3	2.3025	1.1513	0.07875	14.62	0.043
$x_1 * x_2$	0.3275	0.1638	0.07875	2.08	0.285
$x_1 * x_3$	-0.0875	-0.0437	0.07875	-0.56	0.677
$x_2 * x_3$	0.2575	0.1288	0.07875	1.63	0.349

Only 1 term has a small p-value

- Caution: the power (sensitivity) of a t test depends on the residual degrees of freedom.
- With ResDF = 1 the test has low power, and we may have missed some important effects.

Selecting The Important Effects (Cont.)

- In next session we will show that some screening experiments have ResDF = 0.
  - we cannot calculate standard errors for the regression coefficients.
  - so we cannot use t tests.
- By increasing ResDF:
  - we increase the power of t tests.
  - we can calculate a sensible value for R^2.
- Warning: replication does not help with residual plots or with PRESS.
  - these are rarely useful in screening experiments.

In next session we will introduce another method for selecting the important effects

Alternatively we could replicate the design

In This Session We Have…

- Analysed the data from a small screening experiment based on a full factorial 2^3 design.
- Shown that the special features of this design make the multiple regression calculations much simpler.
- Defined and interpreted two new types of plot:
  - interaction plots and main effect plots.
- Explained that although software such as Minitab will produce the usual regression output, R^2, PRESS and residual plots are of little or no use when the residual degrees of freedom (ResDF) are small.
- Explained that t tests will have low power unless the design is replicated.

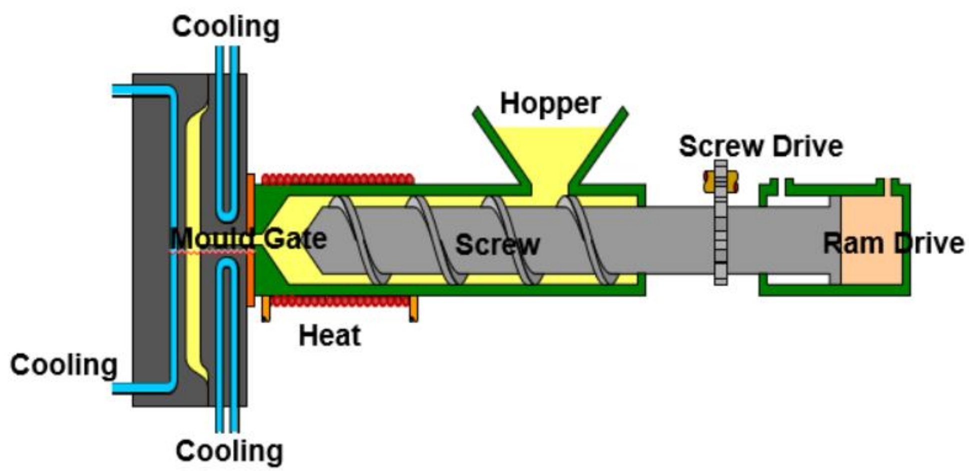
Session 6: Analysing a Two Level Experiment

Tutorial and Exercise

Tutorial

- **Session TS06+07: Analysing a Two Level Experiment**
- **Objectives**

  Analyse a two level screening experiment.
- **Engineering Scenario**
- A study of potential alternative materials for bumper bracket mouldings.
- The brackets are injection moulded (IM).
- In our example, the factors are settings on the IM machine.



Tutorial

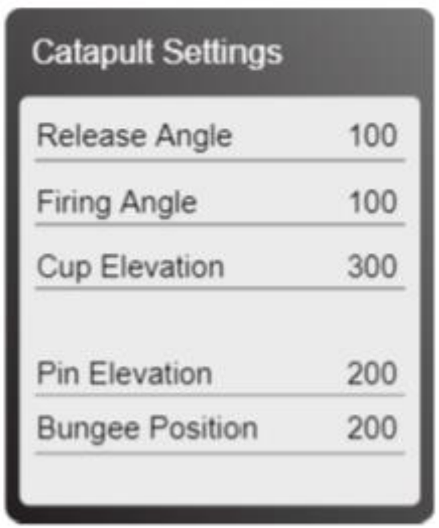
- **Session TS06+07: Analysing a Two Level Experiment**
- **Python Environment**

  A self-guided tutorial has been created as a Colab notebook with pre-designed Python code and notes. For this tutorial, follow the instructions in the notes, upload data files and run the code. No modification of code is required. Interpret the results in accordance with the Technical session. **Stop after Tutorial 06, Task 3.6.**
- **Tutorial Task**
  1. Generate a 2^3 coded design.
  2. Export the design to add the response of the system and then import the edited file.
  3. Generate regression output from the 2^3 design.
  4. Convert the coded design to uncoded units
  5. Examine main effect plots and analyse interaction plots
  6. Examine Plot of the response surface.
  7. Generate and analyse a half normal plot

Exercise

- **Session TS06+07: Analysing a Two Level Experiment**
- **Objective**
- To analyse and discuss the results of an a screening experiment for the Virtual Catapult.

https://sigmazone.com/catapult/



Exercise

- **Session TS06+07: Analysing a Two Level Experiment**
- **Python Environment**

  The exercise has been created as a Colab notebook with notes. For this exercise, follow the instructions in the notes, and create your own code using the accompanying tutorial as a guide. Interpret the results in accordance with the Technical session. **Stop after Exercise 06, Task 5.**
- **Guidelines**

  Use the relevant cells from the Python Colab notebook in Tutorials 06 and 07 to analyse your experiment using regression to fit the appropriate (first order or first order with two-factors interactions) model, and draw conclusions about which factors are significant and you would consider for inclusion in a follow-up 3-level experiment.