Notes:

2k Factorial Experiments - Blocking

Key Learning Points

1. Describe when to use blocking in designed experiments.
2. Describe why blocking is used in designed experiments.
3. Utilize blocking when designing 2k factorial experiments in your improvement project.

What is Blocking?

In the equation Y=f(x1, x2, x3, …) certain X variables may influence the outcome Y, but are not inherently of interest. The addition of blocking variables is an efficient way to account for and "block out" variation introduced by these Xs. Failure to include blocking variables can result in confounding of effects, leading to incorrect conclusions.

DOE Vocabulary

- k1 x k2 x k3 ... Factorial: Description of the basic design. The number of ks is the number of factors. The value of each k is the number of levels of interest for that factor. In 2k experiments, you will always have the value of k1=k2=k3=2. A 23 = 2 x 2 x 2, similarly a 22 = 2 x 2, etc… These types of basic designs indicate that all factors are investigated at only two levels.
- Repetition: Running several experimental runs consecutively using the same treatment combination.
- Replication: Replicating the entire experiment. Replication automatically

implies that you do NOT run several experimental runs consecutively using the same treatment combination.

- Center Points: n repeats or replicates run at the center or midpoint of all quantitative factor levels. Center points can often represent the process parameter levels at their current operating conditions. These points allow you to check the accuracy of a linear fit.
- Curvature: The situation when the output of the process does not act linear at center levels (center points) of all factors.
- Experimental Error: The variation in the data (sum of squares) left over after all significant sources of variability have been accounted for.
- Residual Error: the variability that remains after all the main effects and interactions have been identified.
- Lack-of-Fit: an indication of the fit of a prediction model. If the lack- of- fit p-value is less than your selected a-level, evidence exists that your model does not accurately fit the data.
- Pure Error: The variation in the data (sum of squares) that can only be estimated via true repeat or replicate measurements, either via center points or via repeats/replicate sample size selection.
- Blocking Variable: A factor in an experiment that has undesired influence as a source of variability is called a "block." A block can be a batch of material or a set of conditions likely to produce experimental runs that are more homogenous within the block than between blocks. For example, parts from a single batch of material are likely to be more uniform than parts from different batches. The batch of material would be regarded as the blocking variable.
- Block: Group of homogeneous experimental runs.
- Confounding: One or more effects that cannot unambiguously be attributed to a single factor or interaction.

Adding a Block to 2k Factorials

A factor in an experiment that has undesired influence as a source of variability is called a "block." A block can be a lot of material or a set of conditions likely to produce experimental runs that are more homogenous within the block (or batch) than between blocks.

Suppose you wanted to run a 2 x 2 x 2 factorial. You would like to run the experiment under as homogeneous conditions as possible. Yet you find that two batches of raw material are needed to run the entire experiment.

If you ran the first four runs with batch 1 of raw material and the 2nd four runs with batch 2, you would completely "confound" Factor C. You could not distinguish between the effect because of Factor C and the effect of variation in the raw material.

Notes:

Run	A	B	C	Block
1	-1	-1	-1	I
2	+1	-1	-1	I
3	-1	+1	-1	I
4	+1	+1	-1	I
5	-1	-1	+1	II
6	+1	-1	+1	II
7	-1	+1	+1	II
8	+1	+1	+1	II

We must figure out a way to “spread” the Raw Material effect across the experiment so the differences in Material batches are “seen” by all the main effects. Recall the expanded design matrix for a 23 factorial experiment showing all contrasts.

In general, you make the assumption that “higher order interactions” are not significant (P-Value > 0.05). Here, you can use the contrast for the 3-way interactions to define your blocking variable. The new design would look like:

Run	A	B	C	A*B	A*C	B*C	A*B*C	Block
1	-1	-1	-1	+1	+1	+1	-1	I
2	+1	-1	-1	-1	-1	+1	+1	II
3	-1	+1	-1	-1	+1	-1	+1	II
4	+1	+1	-1	+1	-1	-1	-1	I
5	-1	-1	+1	+1	-1	-1	+1	II
6	+1	-1	+1	-1	+1	-1	-1	I
7	-1	+1	+1	-1	-1	+1	-1	I
8	+1	+1	+1	+1	+1	+1	+1	II

Perform runs 1, 4, 6, and 7 with batch 1 and 2, 3, 5, and 8 with batch 2. This experiment would not allow you to test for the significance of the 3-way interaction, but would allow you to investigate the main effects and 2-way interactions without worrying about confounding these effects with Raw Material batch.

Run	A	B	C	Block
1	-1	-1	-1	I
2	+1	-1	-1	II
3	-1	+1	-1	II
4	+1	+1	-1	I
5	-1	-1	+1	II
6	+1	-1	+1	I
7	-1	+1	+1	I
8	+1	+1	+1	II

Notes:

Run	A	B	C	Block
1	-1	-1	-1	I
2	+1	-1	-1	I
3	-1	+1	-1	I
4	+1	+1	-1	I
5	-1	-1	+1	II
6	+1	-1	+1	II
7	-1	+1	+1	II
8	+1	+1	+1	II

Example

A process engineer in motors is conducting an experiment to improve the armature balancing process.

Step 1: Define the Problem

The Problem:

A process engineer in motors is conducting an experiment to improve the armature balancing process. The response variable is first time pass (FTP), the percent of armatures that are balanced in one operation. The experiment will take 16 runs and requires 100 armatures for each run. Only 8 runs of the experiments can be conducted in one shift.

Step 2: State the Hypotheses

The effect of at least one of the factors on the output is zero.

- H_o: Effect of Cutter Speed = 0
- H_a: Effect of Cutter Speed ≠ 0
- H_o: Effect of Cycle Time = 0
- H_a: Effect of Cycle Time ≠ 0
- H_o: Effect of Incoming Balance = 0
- H_a: Effect of Incoming Balance ≠ 0
- H_o: Effect of Cutter Type = 0
- H_a: Effect of Cutter Type ≠ 0

Step 3: State the Factors and Levels of Interest

For this experiment the factors are:

- Output: Armatures balanced in 1 pass
- Inputs:

Notes:

- Cutter Speed (S)
- Cycle Time (T)
- Incoming Balance (B)
- Cutter Type (C)

Step 4: Create an Appropriate Experimental Data sheet

- **Minitab: Stat > DOE > Factorial > Create Factorial Design**
  - Type of Design: 2-Level Factorial (default generators)
  - Number of factor: 4
  - Design: Full Factorial
  - Number of center points per block: 0
  - Number of replicates for corner points: 1
  - Number of blocks: 2
  - Options: Uncheck randomize runs
  - Factors: S, T, B, and C

	StdOrder	RunOrder	CenterPt	Blocks	S	T	B	C
1	1	1	1	1	1	-1	-1	-1
2	2	2	1	1	-1	1	-1	-1
3	3	3	1	1	-1	-1	1	-1
4	4	4	1	1	1	1	1	-1
5	5	5	1	1	-1	-1	-1	1
6	6	6	1	1	1	1	-1	1
7	7	7	1	1	1	-1	1	1
8	8	8	1	1	-1	1	1	1
9	9	9	1	2	-1	-1	-1	-1
10	10	10	1	2	1	1	-1	-1
11	11	11	1	2	1	-1	1	-1
12	12	12	1	2	-1	1	1	-1
13	13	13	1	2	1	-1	-1	1
14	14	14	1	2	-1	1	-1	1
15	15	15	1	2	-1	-1	1	1
16	16	16	1	2	1	1	1	1

Notes:

Step 5: Run the Experiment and Collect the Data

In real experiments always randomize. For this example standard order has been used.

StdOrder	RunOrder	CenterPt	Blocks	S	T	B	C	FTP
1	1	1	1	1	-1	-1	-1	71
2	2	1	1	-1	1	-1	-1	48
3	3	1	1	-1	-1	1	-1	68
4	4	1	1	1	1	1	-1	65
5	5	1	1	-1	-1	-1	1	43
6	6	1	1	1	1	-1	1	99
7	7	1	1	1	-1	1	1	86
8	8	1	1	-1	1	1	1	70
9	9	1	2	-1	-1	-1	-1	45
10	10	1	2	1	1	-1	-1	65
11	11	1	2	1	-1	1	-1	60
12	12	1	2	-1	1	1	-1	80
13	13	1	2	1	-1	-1	1	100
14	14	1	2	-1	1	-1	1	45
15	15	1	2	-1	-1	1	1	75
16	16	1	2	1	1	1	1	96

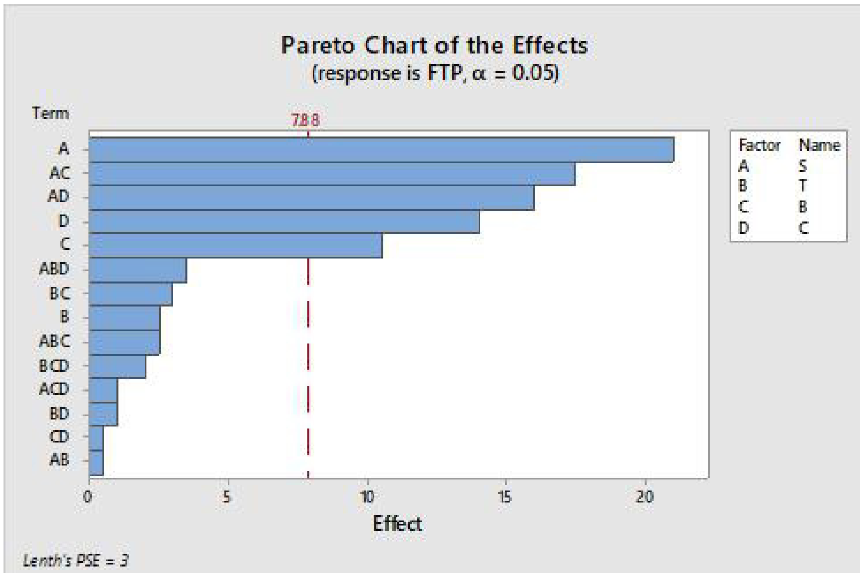
Step 6: Construct The ANOVA Table

- **Minitab: Stat > DOE > Factorial > Analyze Factorial Design**
  - Response: FTP
  - Terms
  - Speed, Time, Balance, Cutter
  - Check: Include Blocks in Model
- Graphs
  - Response: FTP
  - Graphs: Effects Plot
  - Check off Normal and Pareto

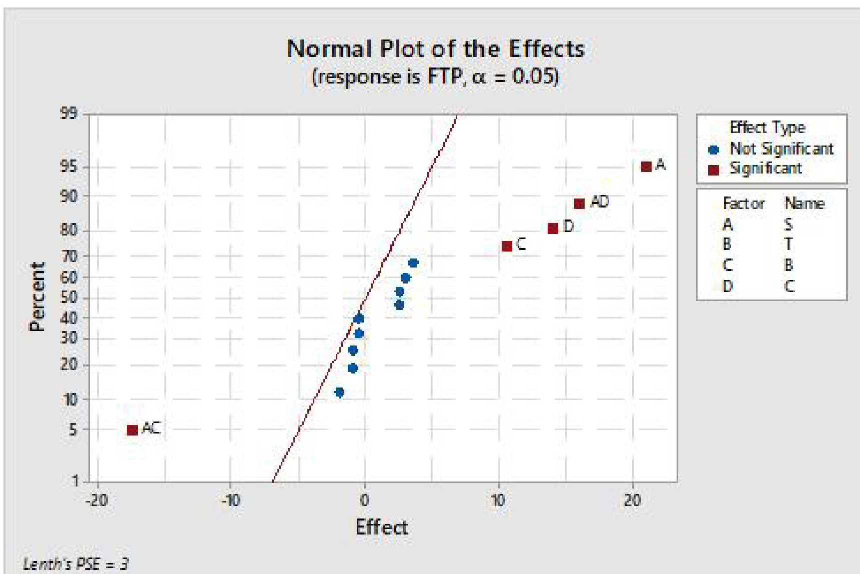
Notes:

Results

Pareto Chart of the Effects



Normap Plot of the Effects



Notes:

Analysis of Variance Table

Notes:

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	15	5415.00	361.00	*	*
Blocks	1	16.00	16.00	*	*
Linear	4	3014.00	753.50	*	*
S	1	1764.00	1764.00	*	*
T	1	25.00	25.00	*	*
B	1	441.00	441.00	*	*
C	1	784.00	784.00	*	*
2-Way Interactions	6	2291.00	381.83	*	*
S*T	1	1.00	1.00	*	*
S*B	1	1225.00	1225.00	*	*
S*C	1	1024.00	1024.00	*	*
T*B	1	36.00	36.00	*	*
T*C	1	4.00	4.00	*	*
B*C	1	1.00	1.00	*	*
3-Way Interactions	4	94.00	23.50	*	*
S*T*B	1	25.00	25.00	*	*
S*T*C	1	49.00	49.00	*	*
S*B*C	1	4.00	4.00	*	*
T*B*C	1	16.00	16.00	*	*
Error	0	*	*		
Total	15	5415.00			

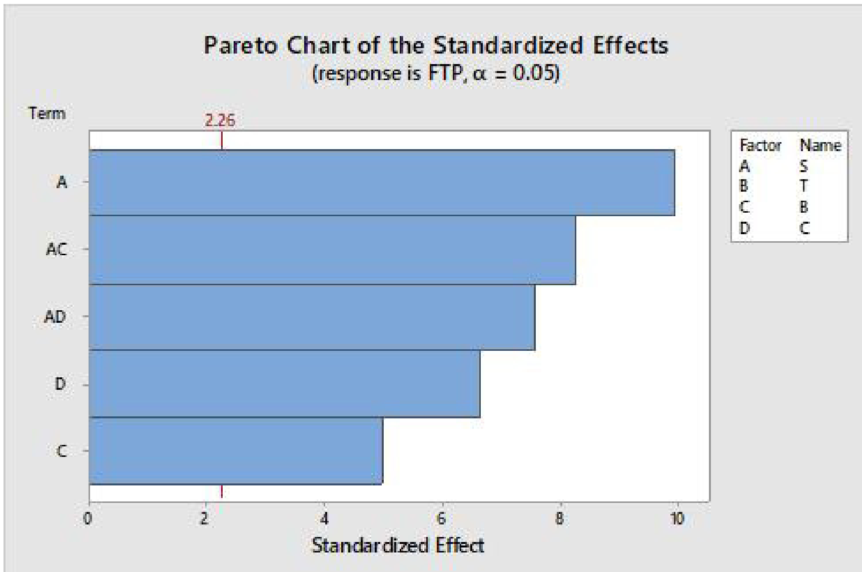
Step 7: Rerun a Reduced Model

Notes:

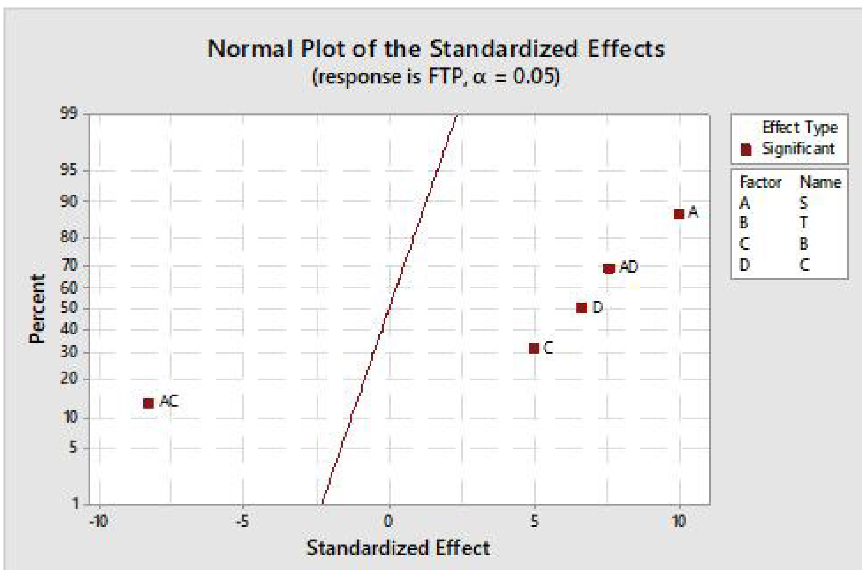
Complete the ANOVA Table again, but remove insignificant effects.

Reduced Model Results

Pareto Chart of the Effects



Normap Plot of the Effects



Analysis of Variance Table

Factorial Regression: FTP versus Blocks, S, B, C

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	6	5254.00	875.67	48.95	0.000
Blocks	1	16.00	16.00	0.89	0.369
Linear	3	2989.00	996.33	55.70	0.000
S	1	1764.00	1764.00	98.61	0.000
B	1	441.00	441.00	24.65	0.001
C	1	784.00	784.00	43.83	0.000
2-Way Interactions	2	2249.00	1124.50	62.86	0.000
S*B	1	1225.00	1225.00	68.48	0.000
S*C	1	1024.00	1024.00	57.24	0.000
Error	9	161.00	17.89		
Total	15	5415.00			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
4.22953	97.03%	95.04%	90.60%

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		69.75	1.06	65.96	0.000	
Blocks						
1		-1.00	1.06	-0.95	0.369	1.00
S	21.00	10.50	1.06	9.93	0.000	1.00
B	10.50	5.25	1.06	4.97	0.001	1.00
C	14.00	7.00	1.06	6.62	0.000	1.00
S*B	-17.50	-8.75	1.06	-8.28	0.000	1.00
S*C	16.00	8.00	1.06	7.57	0.000	1.00

Regression Equation in Uncoded Units

FTP = 69.75 + 10.50 S + 5.25 B + 7.00 C - 8.75 S*B + 8.00 S*C

Equation averaged over blocks.

Alias Structure

Factor	Name
A	S
B	T
C	B
D	C

Aliases

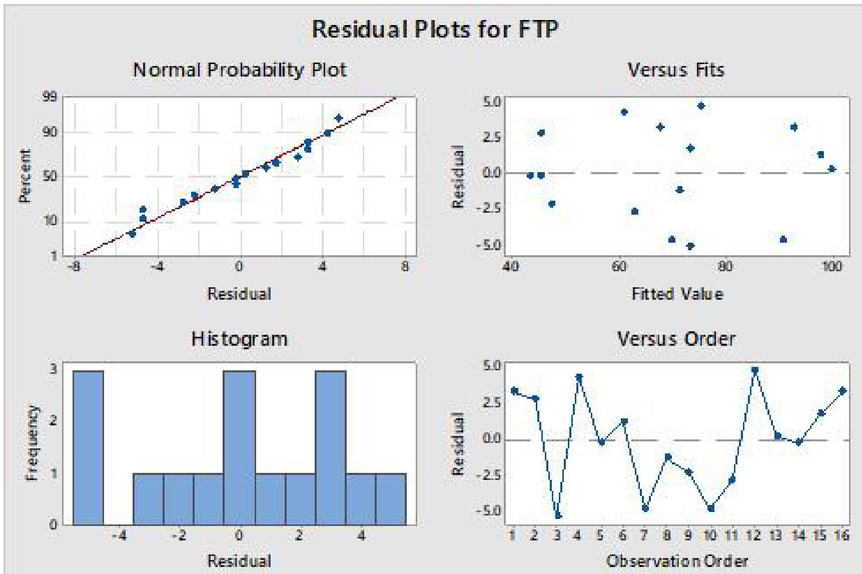
I
Block 1 - ABCD
A
C
D
AC
AD

Effects Plot for FTP

Effects Pareto for FTP

Notes:

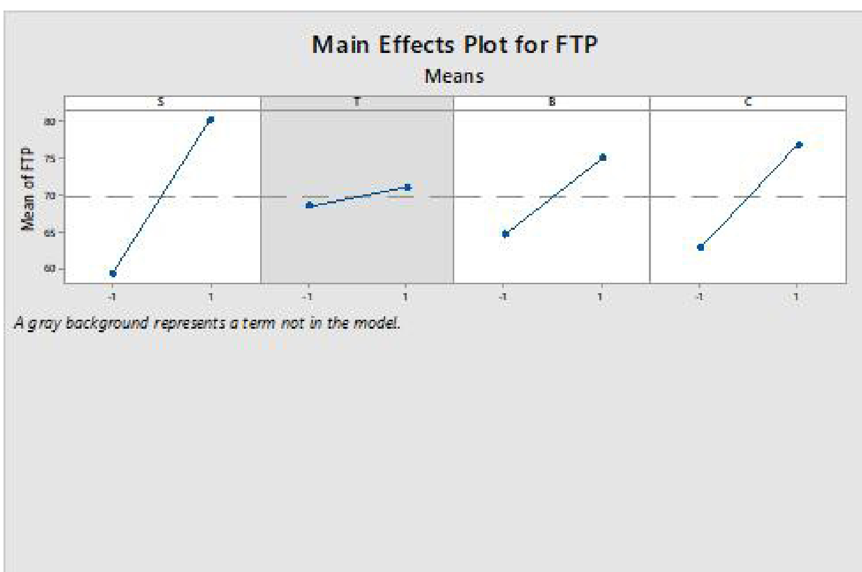
Step 8: Investigate the Residuals Plots to Ensure Model Fit



Step 9: Investigate Significant Main Effects and Interactions

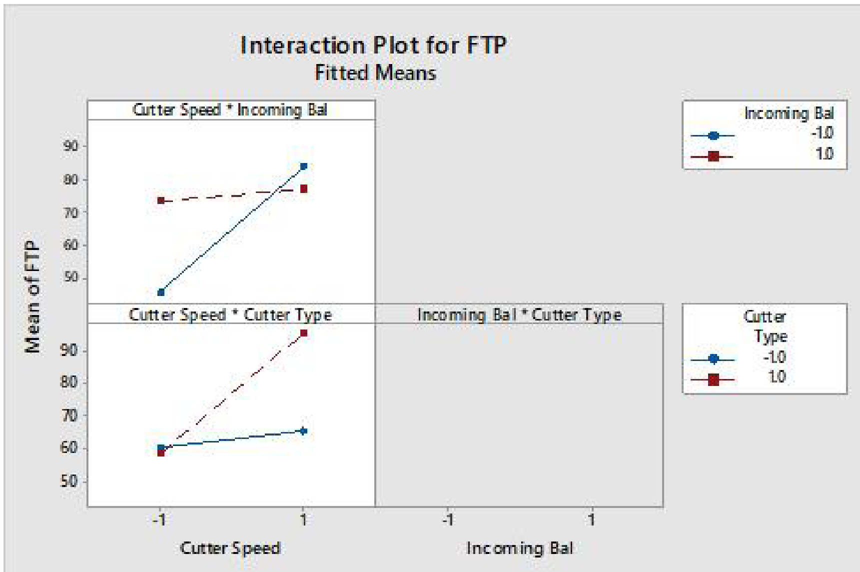
Using the ANOVA table and appropriate graphical tools, investigate the P-Values of the main effects and interactions. Assess the significance of highest order interactions first.

Main Effects Plot

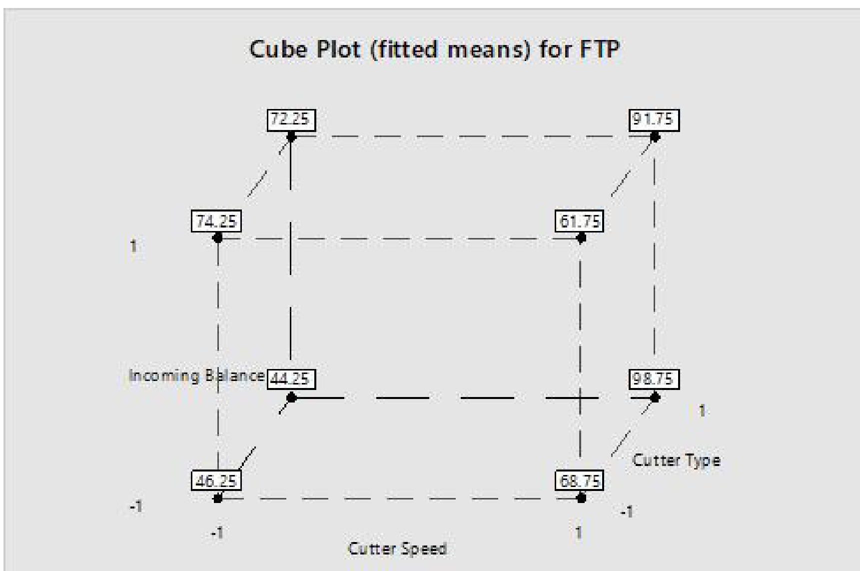


Notes:

Interaction Plot



Cube Plot



Step 10: State the Mathematical Model

Calculate the variation for the main effects and interactions left in the model. State the mathematical model obtained.

Determine the variation that is accounted for by the main effects and interactions left in the model. State the mathematical model obtained. This is quantified by the R^2 value presented in Minitab.

Mathematical Model

Use the coefficients from the analysis to derive the reduced mathematical model (Add the coefficients of the factors.)

Notes:

Notes:

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		69.75	1.06	65.96	0.000	
Blocks						
1		-1.00	1.06	-0.95	0.369	1.00
S	21.00	10.50	1.06	9.93	0.000	1.00
B	10.50	5.25	1.06	4.97	0.001	1.00
C	14.00	7.00	1.06	6.62	0.000	1.00
S*B	-17.50	-8.75	1.06	-8.28	0.000	1.00
S*C	16.00	8.00	1.06	7.57	0.000	1.00

FTP = 69.75 +10.5(S) + 5.25(B) +7(C) - 8.75(S*B) + 8(S*C)

Step 11: Translate the Statistical Conclusion into Process Terms

Practical Question: At what levels should you control the process parameters of cutter speed, cycle time, incoming balance, and cutter type?

Answer:

- Speed = +1
- Time = most cost-effective level
- Balance = -1
- Cutter = +1

Step 12: Replicate Optimum Conditions

Plan the next experiment and/or institutionalize the change.

When Should Blocking Be Used?

Use Blocking in factorial experiments to account for and block out extraneous variation introduced by factors not intended as main experimental factors of interest.

Pitfalls to Avoid

- If blocking is statistically significant, you must account for this variation in future models.
- Use the default generators in Minitab to ensure the design remains balanced.
- Run the experimental runs in random order whenever possible.
- Replicate the experiment for more power. Repeated measures from the

same run do not give the same estimating power as replicates.

- Perform a verifying experiment to prove your best conditions.
- A true optimum may lie outside the design.

Notes: