Taking Advantage of New DOE Tools for Random Block Effects

Martin Bezener, PhD

President & Chief Technology Officer

martin@statease.com

July 2023

1

Making the most of this learning opportunity

Hide the Control Panel

Mute your line

To prevent audio disruptions, all attendees will be muted.

Questions can be posted in the **Question** area. If they are not addressed during the webinar, I will reply via email afterwards.

Questions may also be sent to stathelp@statease.com. Please provide your company name and, if you are using Design-Expert, the serial number (found under Help, About).

Note: The slides and a recording of this webinar will be posted on the Webinars page of the Stat-Ease website within a few days.

Stat-Ease Training:
Sharpen Up Your DOE Skills

***NEW Summer 2023*: ESSENTIALS Series**

ESSENTIALS 1: Foundations of Factorial DOE for Breakthroughs

ESSENTIALS 2: Strategy of Experiments for Accelerating Process Development

ESSENTIALS Series: Response Surface Methods for Peak Performance

ESSENTIALS Series: Optimal Tools for Formulation Development

Each class: 2 half-days of hands-on learning

www.statease.com/training/workshops

Agenda

- Why Block?
- Fixed vs. Random Blocks
- Example
- Conclusion

Agenda

- **Why Block?**
- Fixed vs. Random Blocks
- Example
- Conclusion

Why Block?

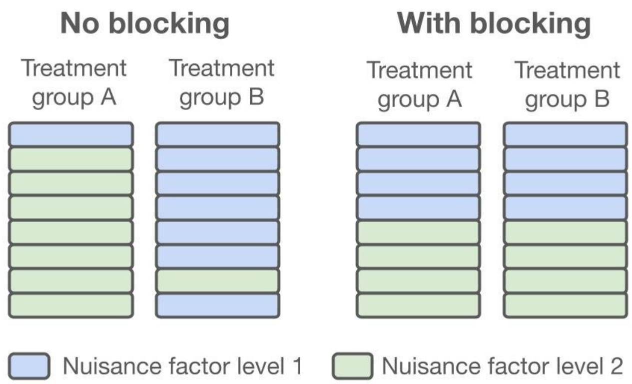
In many experiments, there are external sources of error that may affect the measured responses:

- Variability in different lots of raw material
- Differences in lab sites
- and so on.

We can’t control these and don’t care about their effects, but they do add extra noise that might reduce power and precision in an experiment.

Blocking is one solution – that is, group “similar” experimental units and assign the factor combinations within those groups.

Why Block?



Why Block?

What does blocking ultimately do?

It **removes noise** related to the nuisance factors from the design. This leads to:

- Higher power in factorial designs
- More precision (shorter confidence intervals) in response surface and mixture designs

Blocking is generally done when we have a reason to believe there is a nuisance variable – there isn't a test or algorithm to detect if blocking is necessary.

Agenda

- Why Block?
- **Fixed vs. Random Blocks**
- Example
- Conclusion

Fixed vs. Random Blocks

A **fixed** effects block model looks like:

$$y_{ij} = \mu + \beta_j + \tau_i + \varepsilon_{ij}$$

where μ is the intercept, β_j is the effect from block j, τ_i is the treatment effect of treatment i, and ε_{ij} and is the random error. Each level of the block gets one β coefficient.

A **random** block effects model looks like:

$$y_{ij} = \mu + \tau_i + \sigma_{\beta j} + \varepsilon_{ij}$$

where μ is the intercept, τ_i is the treatment effect of treatment i and ε_{ij} and is the random error. The block variance $\sigma_{\beta j}$ is essentially treated as a random source of error.

Fixed Blocks

Blocks are usually treated as **fixed** effects. Loosely speaking, they are fit and interpreted like the other factors in the model (coefficients), though the inference, p-values, confidence intervals, etc. are either computed slightly differently or ignored.

Fixed blocks are typically reported as a table of coefficients:

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High	VIF
Intercept	81.60	1	1.69	77.77	85.43	
Day 1	-1.92	1				
Day 2	1.92					
A-time	1.03	1	1.11	-1.49	3.54	1.0000
B-temperature	4.04	1	1.11	1.53	6.55	1.0000
C-catalyst	6.20	1	1.11	3.69	8.72	1.0000

Random Blocks

If blocks are fit as random effects, they are essentially treated as a source of variation, just like the experimental error. So instead of one reported error, you'll get two of them.

This method does not produce/fit model coefficients for each level of the block. The individual block effects are analogous to fitted values.

Variance Components

Source	Variance	% of Total	Standard Error	95% CI Low	95% CI High
Block	5.61	24.95	10.46	-14.89	26.11
Residual	16.87	75.05	7.95	7.98	56.23
Total	22.48				

Fixed vs. Random Blocks

Fixed blocks make sense when:

- The blocks can be repeated (e.g. machines)
- There are very few blocks (5 or fewer) in the design

Benefits of fixed blocks:

- Easy to interpret
- Easy to fit - can even do by hand in some situations!

Disadvantages of fixed blocks:

- They use a lot of degrees of freedom. In rare cases, blocks can make things worse if the block effects are small and # of blocks is large (loss of df overpowers the reduction in variance)

Fixed vs. Random Blocks

Random blocks make sense when:

- The blocks are not repeatable (e.g. days) or are a random sample from a huge number of potential blocks (e.g. operators)
- There are many blocks in the design (6 of more)

Benefits of random blocks:

- Simpler model (fewer parameters to estimate)
- Fewer df used for model parameter estimates

Disadvantages of random blocks:

- Not as familiar and popular as fixed blocks
- More complicated to fit

Agenda

- Why Block?
- Fixed vs. Random Blocks
- **Example**
- Conclusion

Example – Shear Strength

This example comes from our tutorial data. Galvanized steel bars are bonded, and the shear strength of the bonded bars is measured.

The factors are:

- **Temperature:** 375, 400, 450 degrees F
- **Time:** 30, 35, 40 seconds

The experiment was repeated over 12 days. A total of 118 independent runs were performed, roughly 10 per day.

Because the days are not repeatable and they are simply a random sampling of all possible days, blocks are treated as random effects in the analysis.

The Design

Block	Run	Factor 1 A:Temperature °F	Factor 2 B:Time sec	Response 1 Shear Strength psi
July 11	1	375	30	1226
July 11	2	400	30	1898
July 11	3	450	30	2142
July 11	4	375	35	1472
July 11	5	400	35	2010
July 11	6	400	35	1882
July 11	7	400	35	1915
July 11	8	400	35	2106
July 11	9	450	35	2352
July 11	10	375	40	1491
July 11	11	400	40	2078
July 11	12	450	40	2531
July 16	13	375	30	1075
July 16	14	400	30	1790
July 16	15	450	30	1843
July 16	16	375	35	1121
July 16	17	400	35	2175
July 16	18	450	35	2274
July 16	19	375	40	1691
July 16	20	400	40	2513
July 16	21	450	40	2588
July 20	22	375	30	1172

Fixed vs. Random Analysis

Variance Components

Source	Variance	% of Total	Standard Error	95% CI Low	95% CI High
Block	3480.71	21.68	1985.94	-411.67	7373.08
Residual	12571.41	78.32	1763.80	9721.67	16894.72
Total	16052.12				

Coefficients in Terms of Coded Factors

Factor	Coefficient Estimate	df	Standard Error	95% CI Low	95% CI High
Intercept	2375.19	1	22.71	2330.14	2420.23
July 11	-103.47	11			
July 16	-79.12				
July 20	-8.46				
Aug. 7	117.03				
Aug. 8	9.21				
Aug. 14	3.88				
Aug. 2	38.99				
Aug. 22	60.88				
Sep. 11	-100.19				
Sep. 24	58.43				
Oct. 3	12.17				
Oct. 1	-9.35				

Agenda

- Why Block?
- Fixed vs. Random Blocks
- Example
- **Conclusion**

Conclusion

Tips and Tricks:

- If your design has 6 or fewer levels, use fixed blocks (even if they are truly random) because of numerical fitting issues.
- Decided on fixed vs random BEFORE designing and analyzing the experiment – don't choose the one that "looks better".
- Don't go crazy with blocking – it's often helpful but isn't a total free pass. In severe cases blocking won't help and may actually make things worse.
- Consider investing some time into learning about random effect blocks – it will serve you well into the future!

StatEase
statistics made easy

Make the most from every experiment!℠

Thanks for listening!

martin@statease.com