StatEase
statistics made easy
Make the most from every experiment!

Keys to Building the Perfect Response Surface Design

Shari Kraber, MS Applied Statistics
stathelp@statease.com

November 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 Stat-Ease YouTube channel and the Webinars page of the Stat-Ease website within a few days.

2

Stat-Ease, Inc.

Design of Experiments Software and Training

www.statease.com

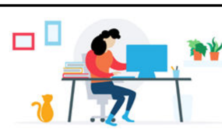
StatEase 360

DESIGN EXPERT

YouTube Statistics Made Easy by Stat-Ease

@StatisticsMadeEasybyStatEase

Stat-Ease Training: Sharpen Up Your DOE Skills



Modern DOE for Process Optimization

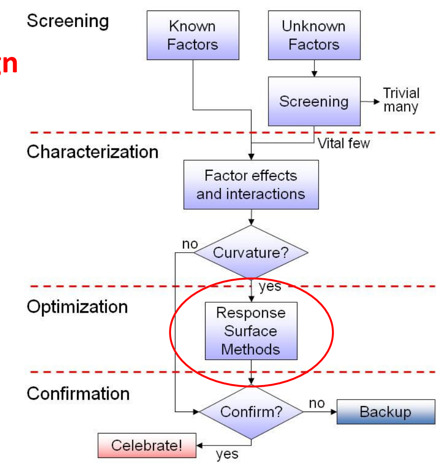
Mixture Design for Optimal Formulations

Individuals	Teams (6+ people)
Improve your DOE skills	Choose your date & time
Topics applicable to both novice and advanced practitioners	Add company case studies

Email list sign up:
www.statease.com/publications/signup/

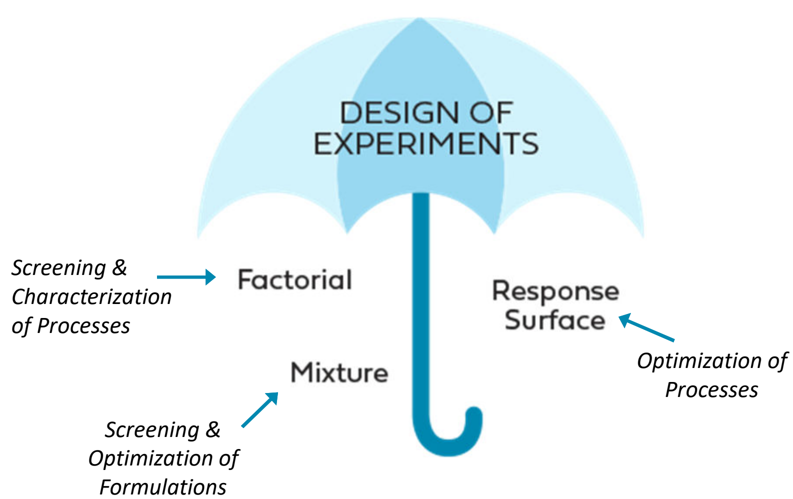
Agenda

- **Fundamentals of an RSM design**
- Central Composite design
- Box-Behnken design
- Optimal design
- Wrap Up



DOE Umbrella

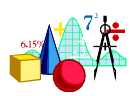
Where do Response Surface (RSM) designs fit?



Key Properties for an RSM Design

1. **Fit desired polynomial** – model points are selected to estimate the polynomial chosen by the experimenter. (Quadratic usually)
2. **Lack of fit test** – additional information to test model adequacy.
   - ☑Include 4-5 more unique design points than coefficients in the model.
   - ☑Include 4-5 replicates to provide an estimate of "pure" error.
3. **Low prediction error** - generate useful predictions throughout the region of interest.
4. **Reasonable number of runs** – just enough runs to get necessary information.

1. Polynomial Models



Most response surface designs fit a full quadratic model:

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2$$

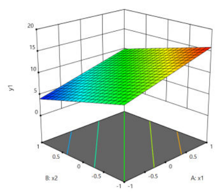
Shape parameters *(pictures on following slide)*:

- Intercept – a horizontal plane.
- Linear terms – slopes (gradients) of the plane.
- Two-Factor interactions – twists in the plane.
- Squared terms – symmetric curvature.
- Cubic terms – asymmetry (inflection).

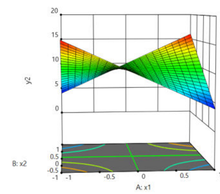
1. Polynomial Models

Coefficients define shape of response surface

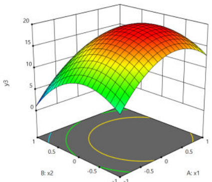
Linear terms



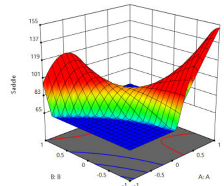
Two-Factor Interaction



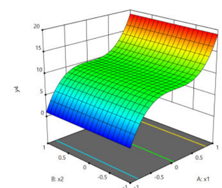
Quadratic



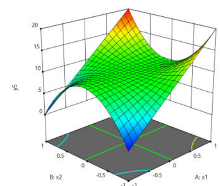
Quadratic



Cubic



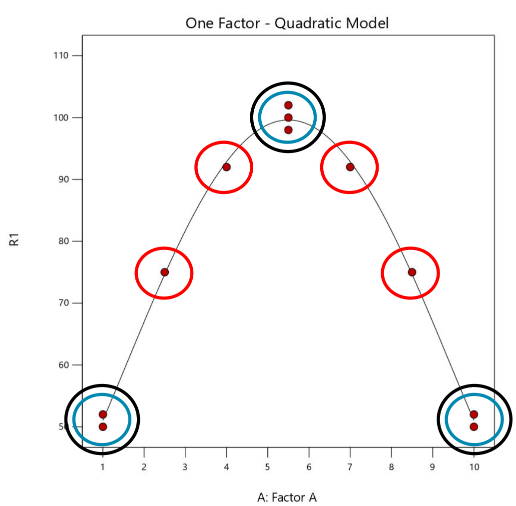
Cubic



2. Selecting Design Points

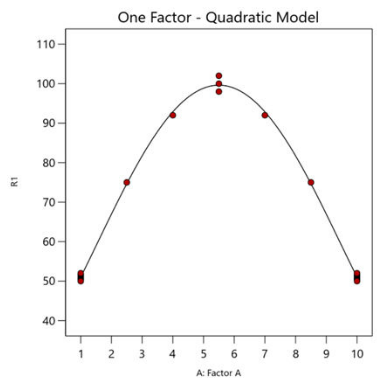
Model, Lack of Fit, Replicates

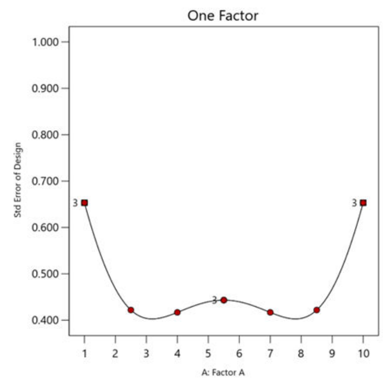
- ✓ Model points
- ✓ Lack of Fit points
- ✓ Replicates



3. Prediction Error

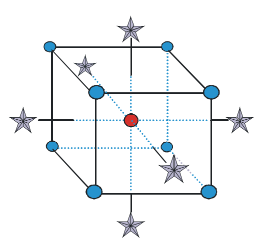
Design point locations should keep the standard error of predictions in the middle of the design relatively low.

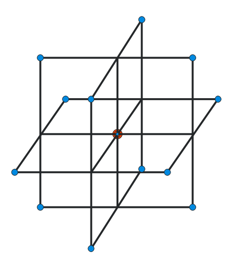




Top Three Response Surface Designs

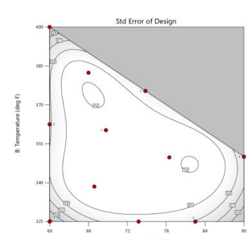
Central Composite





Box-Behnken

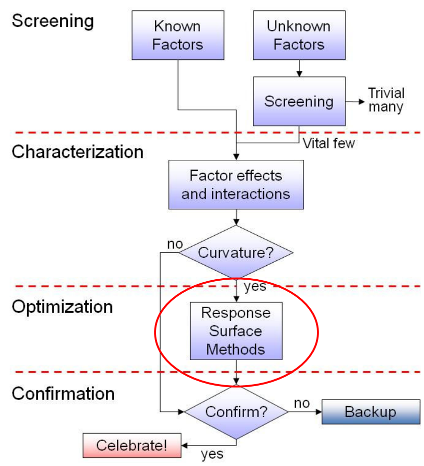
Optimal



All of these designs are statistically sound, and likely more efficient than what an experimenter would create by hand.

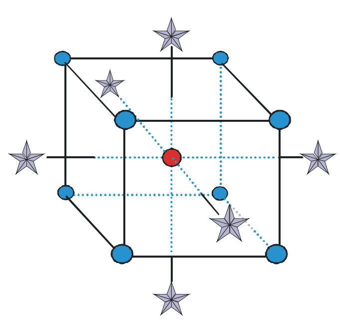
Agenda

- Fundamentals of an RSM design
- **Central Composite design**
- Box-Behnken design
- Optimal design
- Wrap Up



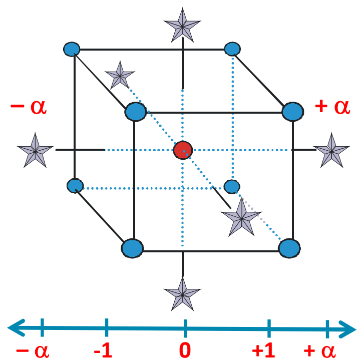
Central Composite Design

Model points



- Two-level factorial points
  - Estimate linear effects and two-factor interactions.
- Center points
  - Estimate quadratic effects, replicated to estimate pure error and tie blocks together.
- Star (axial) points
  - Estimate pure quadratic effects.

Central Composite Design
Template for 3 Factors



α (alpha) is the coded distance from the center to the axial (star) points.

	A	B	C
Factorial points:	−1	−1	−1
	1	−1	−1
	−1	1	−1
	1	1	−1
	−1	−1	1
	1	−1	1
	−1	1	1
	1	1	1
Axial (star) points:	−α	0	0
	α	0	0
	0	−α	0
	0	α	0
	0	0	−α
	0	0	α
Center points:	0	0	0
	0	0	0
	0	0	0
	0	0	0
	0	0	0
	0	0	0

Central Composite Design
Example: 3-Factor design – Tutorials “Conversion”

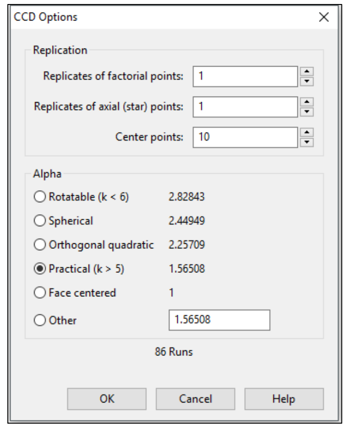
Std	Block	Space Type	Factor 1 A:Time min.	Factor 2 B:Temperature deg C	Factor 3 C:Catalyst %
1	Day 1	Factorial	40	80	2
2	Day 1	Factorial	50	80	2
3	Day 1	Factorial	40	90	2
4	Day 1	Factorial	50	90	2
5	Day 1	Factorial	40	80	3
6	Day 1	Factorial	50	80	3
7	Day 1	Factorial	40	90	3
8	Day 1	Factorial	50	90	3
9	Day 1	Center	45	85	2.5
10	Day 1	Center	45	85	2.5
11	Day 1	Center	45	85	2.5
12	Day 1	Center	45	85	2.5
13	Day 2	Axial	36.591	85	2.5
14	Day 2	Axial	53.409	85	2.5
15	Day 2	Axial	45	76.591	2.5
16	Day 2	Axial	45	93.409	2.5
17	Day 2	Axial	45	85	1.6591
18	Day 2	Axial	45	85	3.3409
19	Day 2	Center	45	85	2.5
20	Day 2	Center	45	85	2.5

Customize the axial runs by rounding or modifying.

Choices for Alpha
Watch Out for Axial Points Going Too Far Out

CCD Options (highlights):

k = 6 factors



- ❑ **Rotatable*** *(default k< 6 factors)*
  Ideal statistically but increases quickly as the number of factors (k) goes up.
- ❑ **Practical** *(default k > 5 factors)*
  Alpha = $k^{1/4}$ (4^{th} root) pushes axials outside the factorial box, but not too far.
- ❑ **Face centered** *(not advised for k>8)*
  Alpha = 1 makes CCDs only 3 level, but variance inflation factors (VIFs) increase with more factors.

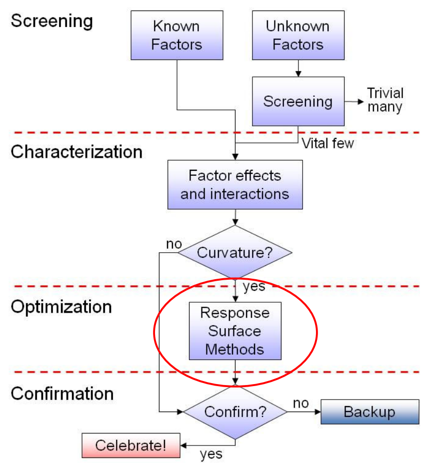
*keeps prediction error equal at points equidistant from the center

Central Composite Design
Keys for usage

- ✓ **Axial points:** watch the experimental values of the axial points
  - Use practical alpha if needed
  - Manually adjust the factor setting if a point is too far out (i.e., goes negative)
- ✓ **Center point replicates:** use the default number of replicates for the center point – keeps the prediction error in the middle of the design low
- ✓ **Blocking:** use statistical blocking if the design needs to be run over multiple days, or to remove another potential source of noise (details beyond this webinar)

Agenda

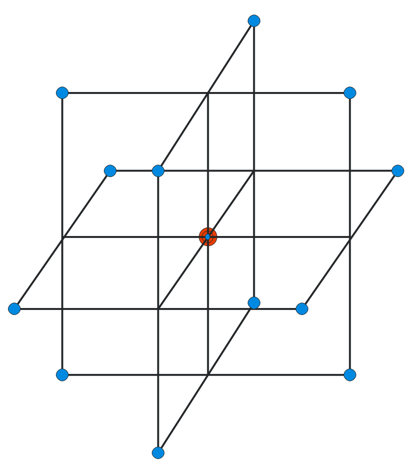
- Fundamentals of an RSM design
- Central Composite design
- **Box-Behnken design**
- Optimal design
- Wrap Up



Box-Behnken Design

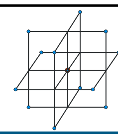
Point Layout *(3-factor example)*

The geometry of the 3-factor design involves 12 points lying on a **sphere** about the center (in this case at √2) with 5 replicates of the center point.



Box-Behnken Design

Design Matrix *(3-factor example)*

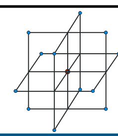


No runs where all factors are set to the extremes.

A	B	C
−1	−1	0
+1	−1	0
−1	+1	0
+1	+1	0
−1	0	−1
+1	0	−1
−1	0	+1
+1	0	+1
0	−1	−1
0	+1	−1
0	−1	+1
0	+1	+1
*0	0	0

**We suggest 5 center points*

Box-Behnken Design



Description:

- Efficient three-level design for fitting second-order (quadratic) response surfaces for up to 21 factors.

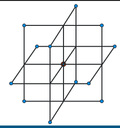
Advantages:

- Only 3 levels (vs 5 for CCD)
- Can be blocked orthogonally (except for k = 3 factors)
- Rotatable (for k= 4,7) or nearly so

Drawbacks:

- Not as flexible as CCD which allows:
  1. Factorial with center points (stop here if no curvature)
  2. Second block of axial (star) points only if needed

Box-Behnken discussion



Quote from client A:

"I use the Box-Behnken design because it does not have points at the corners of the design space, so I don't have to worry about those high-high-high or low-low-low factor combinations that potentially don't work for my process."

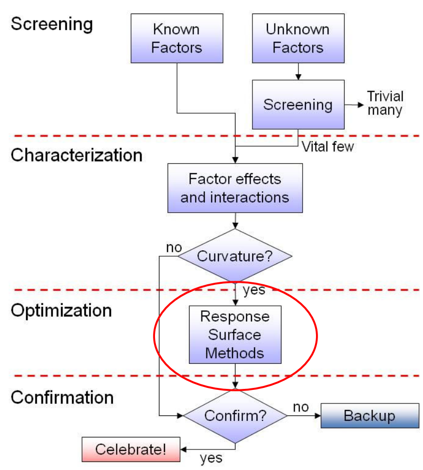
Quote from client B:

"I never use the Box-Behnken because it does not include points at those extreme combinations of the design space. The Central Composite design fills my process space better."

Choose the design that fits the problem!!

Agenda

- Fundamentals of an RSM design
- Central Composite design
- Box-Behnken design
- **Optimal design**
- Wrap Up



Optimal (custom) Design

Description:

- Computer-generated design to fit a selected model. Include lack of fit points and replicate points for robustness (defaults).

Advantages:

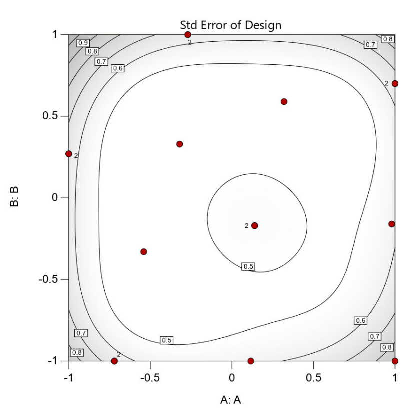
- Good design properties if software defaults used
- Customize the factor types (numeric/discrete/categoric)
- Customize the polynomial model
- Add constraints to fit an irregular-shaped design space

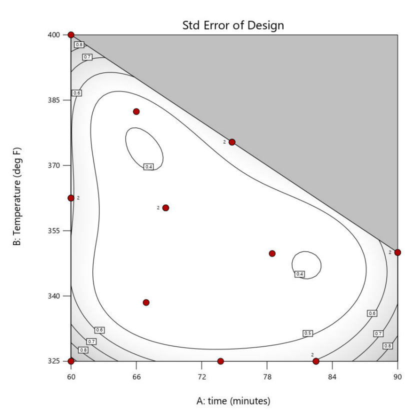
Drawbacks:

- Different point layout each time design generated
- You don't control the number of levels (unless discrete factors)
- Points may be at inconvenient values (may need to round)

Optimal Design

Examples – 2 factor, quadratic model





Optimal Design
Keys to Building

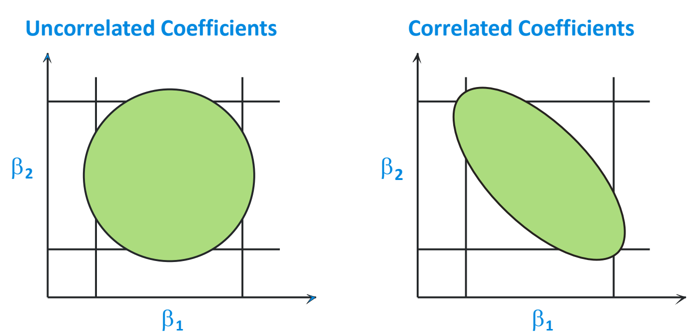
1. Select **polynomial** - that will likely approximate the actual response surface. (Refer to the polynomial pictures – which one most likely?)
   - *Default is a quadratic model.*
2. Select the **optimality** criteria (see next slides).
   - *Default is I-optimal for response surface design.*
3. Select **point selection** method (see next slides).
   - *Default is Both Exchanges, with several designs built using point exchange and several built using coordinate exchange. Most optimal design is displayed to the experimenter.*

Recommendation: Use the software defaults!!!

Optimality: D-optimal

Statistical detail

Goal: D-optimal design minimizes the determinant of the $(X'X)^{-1}$ matrix. This minimizes the volume of the confidence ellipsoid for the coefficients and **maximizes information about the polynomial coefficients**.

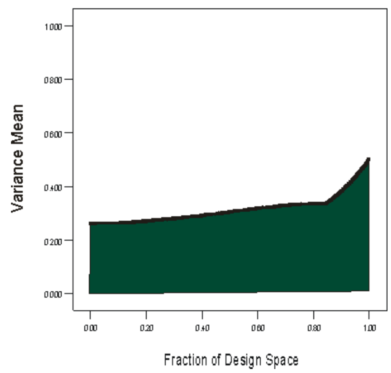


Take away: Best precision of coefficients

Optimality: I-optimal

Statistical detail

An I-optimal design seeks to minimize the integral of the prediction variance across the design space. These designs are built algorithmically to provide ***lower*** *integrated* ***prediction variance across the design space.*** This equates to minimizing the area under the fraction of design space (FDS) curve.



Take away: Lowest standard error of predictions

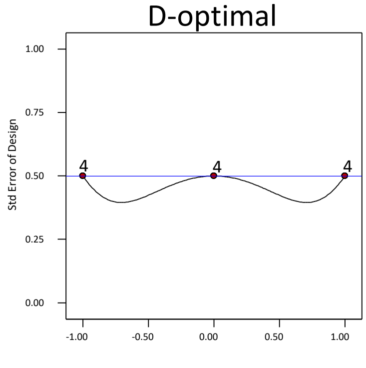
D-optimal versus I-optimal

One Factor 12 run Design

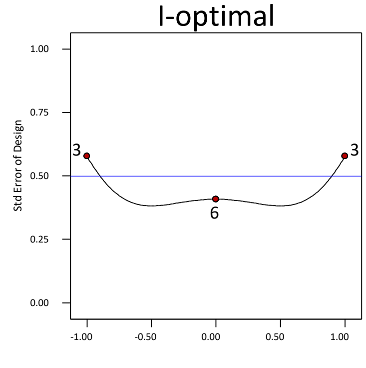
Statistical detail

One factor twelve run quadratic model designs:

D-optimal



I-optimal



Optimality Choice

What you need to remember:

- D-optimal is precise estimation of model coefficients

  Best for factorial designs that detect effects.

- I-optimal is precise estimation of the response

  Best for response surface designs that optimize predictions.

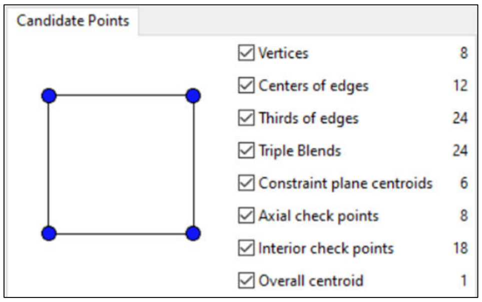
These are the Stat-Ease software default settings.

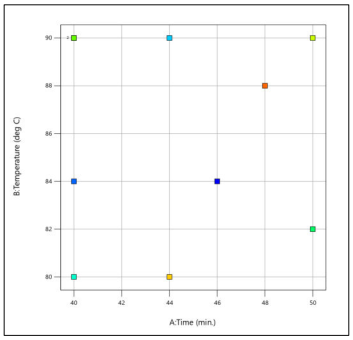
Design Point Selection algorithm

Statistical detail

Start with an initial selection of design points and then switch points in and out, keeping the set with the highest **optimality** score. Points are selected and switched 1 at a time, then 2 at a time, etc., based on either:

Point Exchange or Coordinate Exchange





Design Point Selection
Both Exchange Algorithm

1. To increase the chances of finding the global optimum, a number of designs are built. Using the "both exchange" option:
   - Half the designs are built using point exchange.
   - Half the designs are built using coordinate exchange.
2. The most optimal of these builds is selected as the final design.
3. "**Both Exchanges**" is our recommended point selection algorithm.

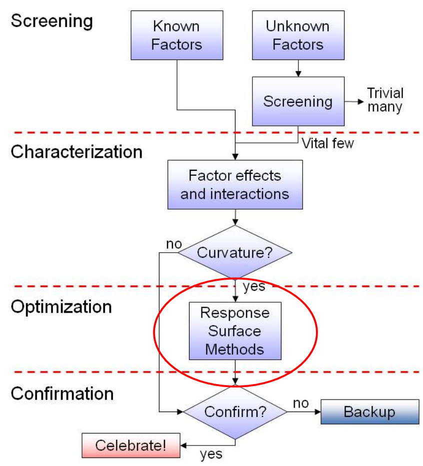
Optimal (custom) Design Summary

- ✓ Flexible design options
- ✓ Factors can be
  - numeric – continuous
  - numeric – discrete (1, 4, 7)
  - categoric – classes/names
- ✓ Multilinear constraints can be added to cut off a portion of the design space that would not provide measurable results.
- ✓ Polynomial models – linear, quadratic, cubic, custom, etc.

While "flexibility" is great, follow good statistical design principles!

Agenda

- Fundamentals of an RSM design
- Central Composite design
- Box-Behnken design
- Optimal design
- **Wrap Up**



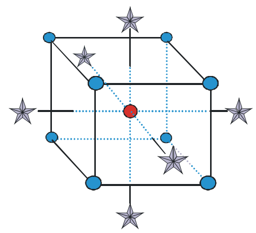
Choosing a Response Surface Design

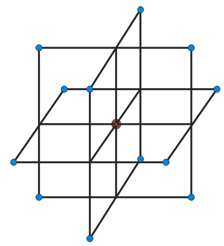
Criteria	Central Composite	Box-Behnken	Optimal
Factor levels	5, 3 optional	3	varies
Design Space	Axial points located outside factorial box	Factorial cube doesn't include corners	Defined by factor limits, not limited to cube
Includes Lack of Fit	Yes	Yes	With defaults
Includes Replicates	Yes – center point	Yes – center point	With defaults, chosen by optimality
Low Prediction Error	Yes	Yes	With defaults
Choose when:	Region of interest is a cube	Region of interest is a sphere	Region of interest is custom

Final Notes

All of these designs are statistically sound, and likely more efficient than what an experimenter would create by hand.

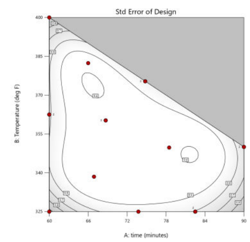
Central Composite





Box-Behnken

Optimal



Choose the design that fits the problem!

Continuing Education

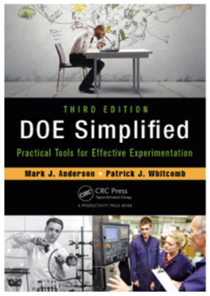
Stat-Ease Webinars: www.statease.com/webinars/

Keys to Analyzing Response Surface Designs – November 30

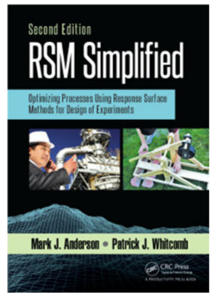
The follow-up to today's webinar dives into the analysis of RSM designs. Learn more about the automated model selection methods and the keys that confirm a good analysis.

Resources*

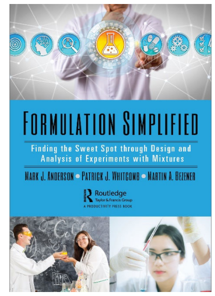
3rd edition 2015

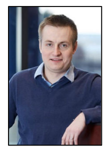


2nd edition 2016



1st edition 2018





** Taylor & Francis/CRC/ Productivity Press New York, NY.*

Self-study options for learning more

YouTube Channel: www.youtube.com/c/StatisticsMadeEasybyStatEase

Playlist: New to DOE?

A collection of webinars on basic to intermediate-level topics.

Stat-Ease Academy: www.statease.com/training/academy/

Self-paced online courses covering the basics of factorial and fractional-factorial designs.

StatEase
statistics made easy

*Make the most from every experiment!*SM

Thank you for listening!

Questions? Email stathelp@statease.com