CARLSON
SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

Exploring Best Practices in the Design of Experiments

Chris Nachtsheim
University of Minnesota
Carlson School of Management
and
School of Statistics

Joint work with Brad Jones
SAS/JMP

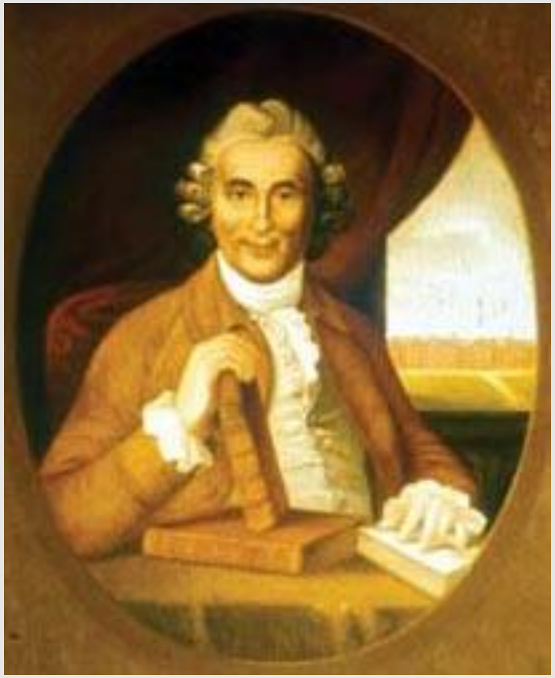
Overview

- Where have we been?
- Where are we now?
- Emerging paradigm: problem-driven design
- Six Examples
- What's driving the software?
- Definitive screening and an example
- The future?

Where have we been?

Where have we been?

- James Lind, 1753: "A Treatise on Scurvy"
- First (published) one-way layout



From "A Treatise…"

"On the 20th May, 1747, I took twelve patients in the scurvy on board the Salisbury at sea. Their cases were as similar as I could have them. They all in general had putrid gums, the spots and lassitude, with weakness of their knees. …

- Two of these were ordered each a quart of cyder a day…
- Two others took twenty five gutts of elixir vitriol three times a day upon an empty stomach…
- Two others took two spoonfuls of vinegar three times a day upon an empty stomach…
- Two of the worst patients, with the tendons in the ham rigid (a symptom none the rest had) were put under a course of sea water…
- Two others had each two oranges and one lemon given them every day…
- The two remaining patients took the bigness of a nutmeg three times a day…

The consequence was that the most sudden and visible good effects were perceived from the use of the oranges and lemons"

Where have we been?

- **Gergonne: 1815**
- **Designs for polynomial regression, response surface design**

- **S. C. Peirce: 1870s**
- **Randomization**

- **K. Smith, 1918: *Biometrika*, Optimal design for polynomial regression**

R. A. Fisher put it all together

Fisher, 1920s:

- Randomization as mathematical basis for analysis
- Local control and blocking
- Replication
- Factorial designs
- Split plot designs
- Confounding
- ANOVA
- F, t distributions
- Etc., etc.

1920's-1950's: Orthogonality is the driving principle

- Fisher, Yates: need for ease of computation, independence of effects

- R. C. Bose, C.R. Rao, and Indian School: Combinatorics, BIBDs, PBIBDs

- Finney, 1945: Fractional replication

- Plackett and Burman, 1946

…culminating in the 2^{k-p} System

Vol. 3, No. 3 TECHNOMETRICS August, 1961

The 2^{k-p} Fractional Factorial Designs*
Part I.

G. E. P. Box and J. S. Hunter

Statistics Department, University of Wisconsin and Mathematics Research Center, University of Wisconsin

1950s: Baby steps away from orthogonality

Journal of the Royal Statistical Society
SERIES B (METHODOLOGICAL)
Vol. XIII, No. 1, 1951

ON THE EXPERIMENTAL ATTAINMENT OF OPTIMUM CONDITIONS

By G. E. P. Box and K. B. Wilson
Imperial Chemical Industries, Dyestuffs Division Headq
Blackley, Manchester

A BASIS FOR THE SELECTION OF A RESPONSE SURFACE DESIGN*

G. E. P. Box
Princeton University
AND
NORMAN R. DRAPER
University of North Carolina and Imperial Chemical Industries

Also Box and Lucas, 1959, Nonlinear Design

Optimal Design—Pioneers

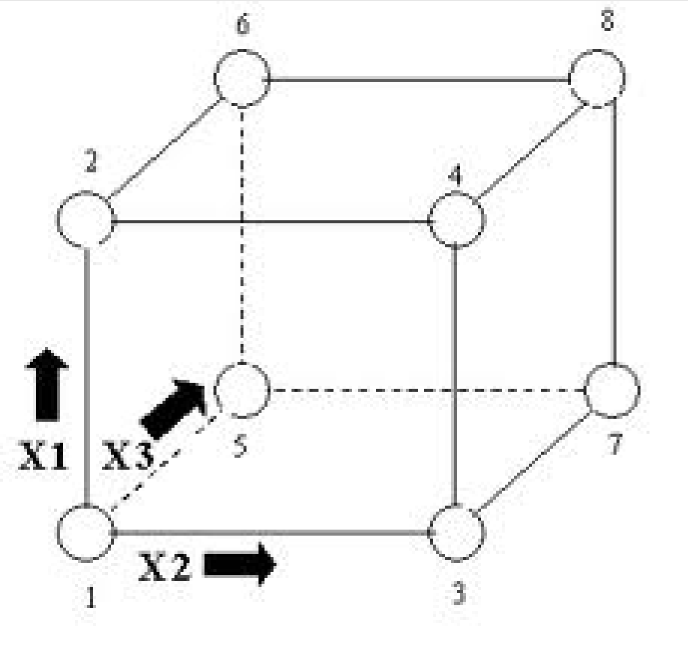
- Kiefer and Wolfowitz (1959)
- Fedorov (1969)
- Wynn (1969)

Shift to problem-driven design (70s, 80s 90s, 00s)

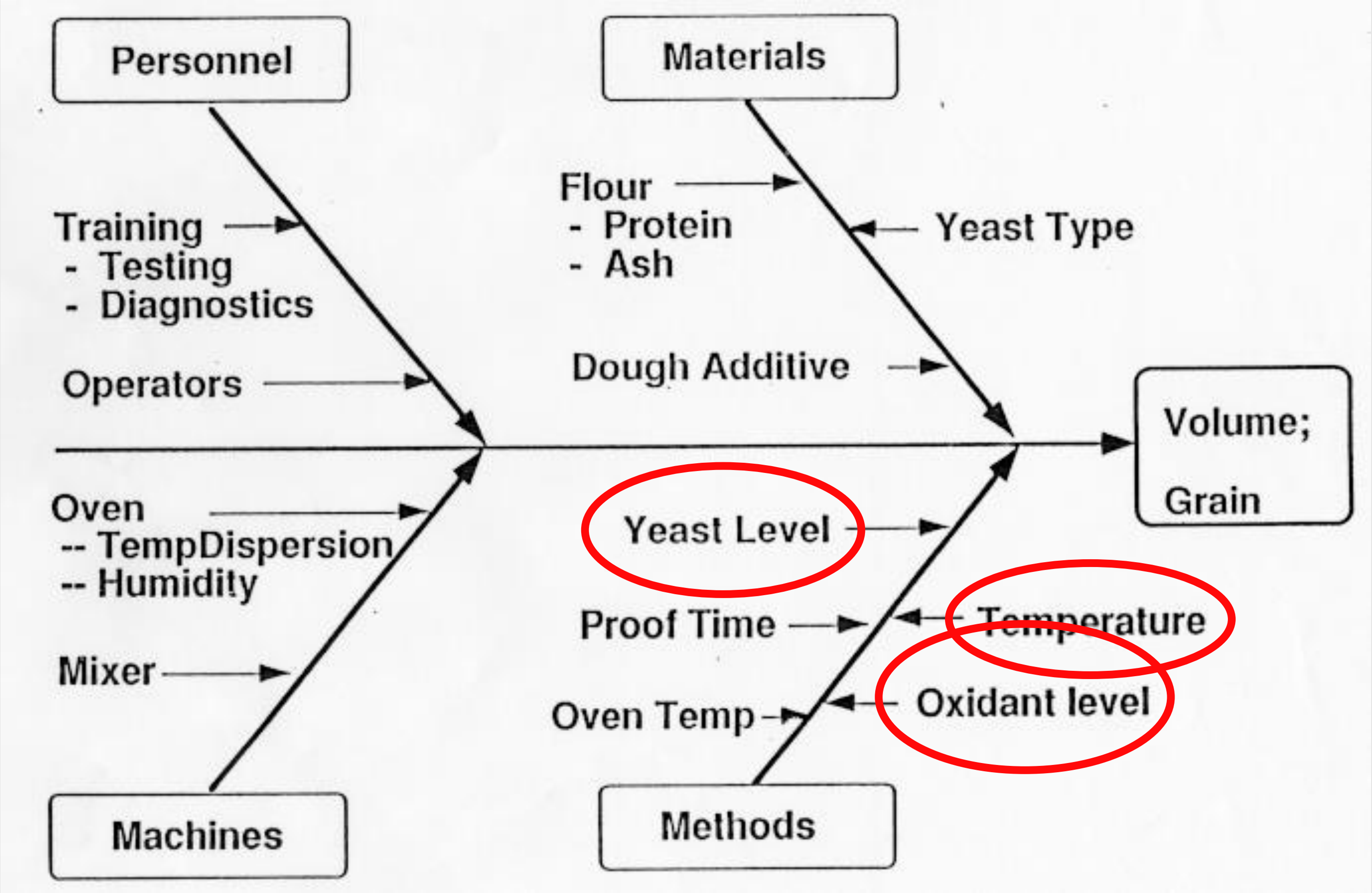

- Toby Mitchell---Detmax
- Constrained mixture designs—John Cornell, Ron Snee, Greg Piepel
- Design augmentation
- Optimal blocking (Cook, Nachtsheim ’87, Donev, Atkinson ’87)
- Model robust design (Lauter, 1974)
- Bayesian design (Chaloner, ‘80, DuMouchel and Jones,’ 94
- Optimal split-plot designs, Goos, ‘2002


Where are we now?

What is a Factorial Experiment?



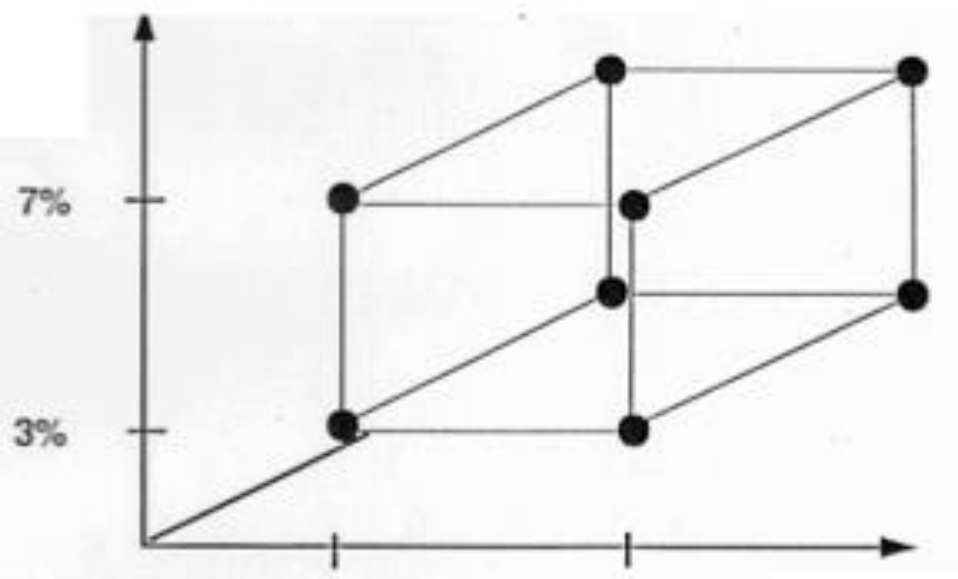
Bread Baking Optimization: Which Factors?



Factorial Design Approach

One Test at Each Corner (n=8)
(2x2x2 factorial design, 3 factors, 2 levels)

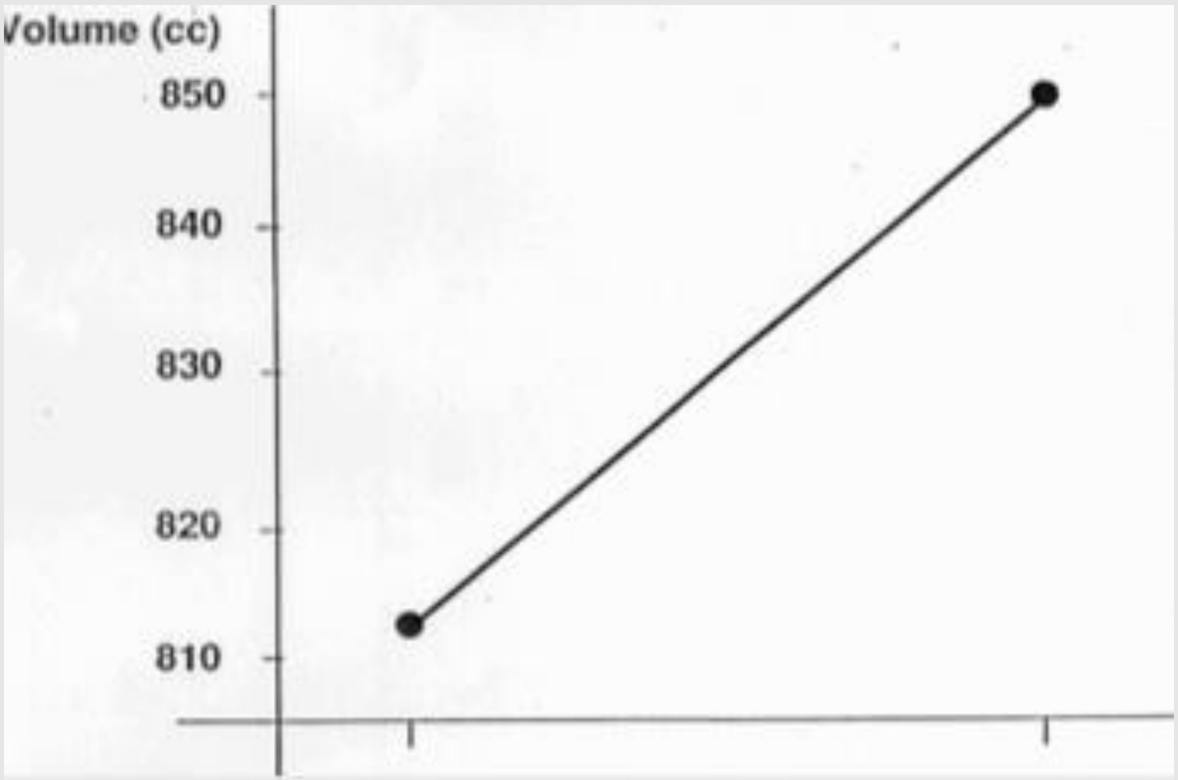
Yeast Level



62 94

Proof Temperature

Analysis: Main Effects Plot

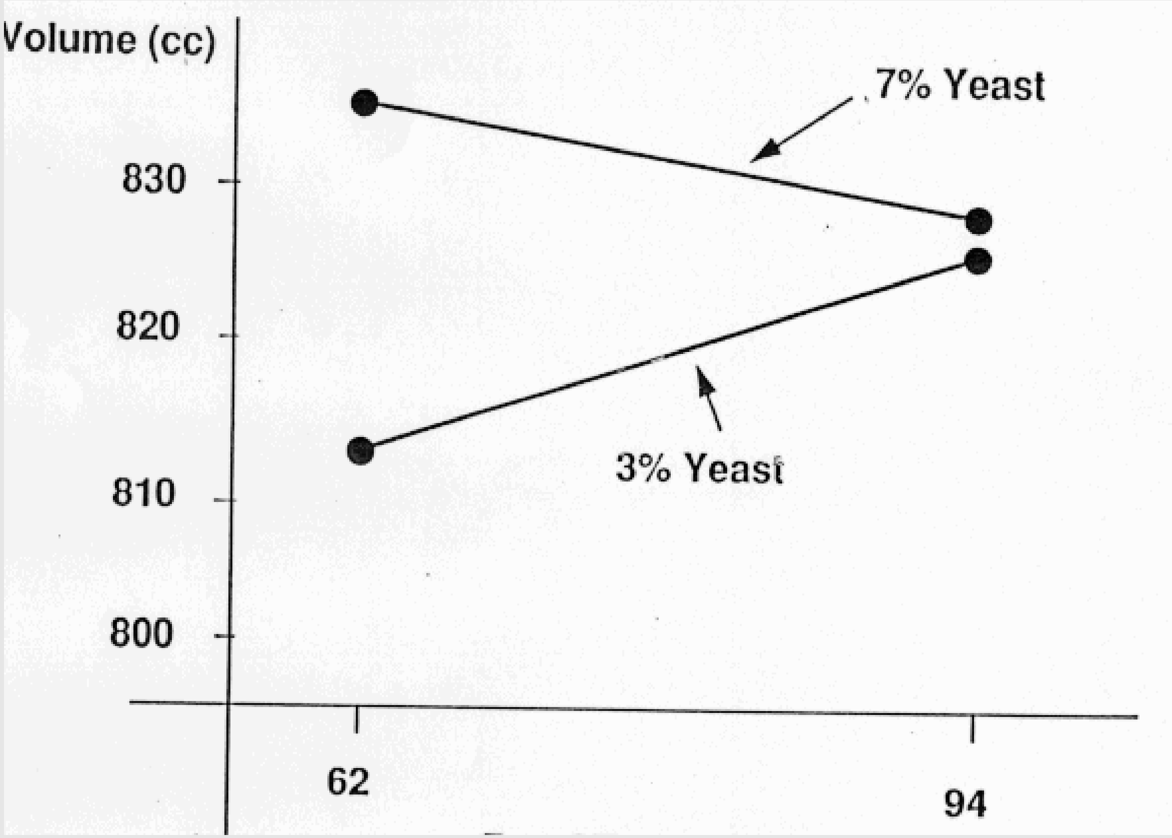


42.2

77.8

Oxidant

Bonus: We can identify interactions



Proof Temperature

JMP Demo: Bread Baking Experiment

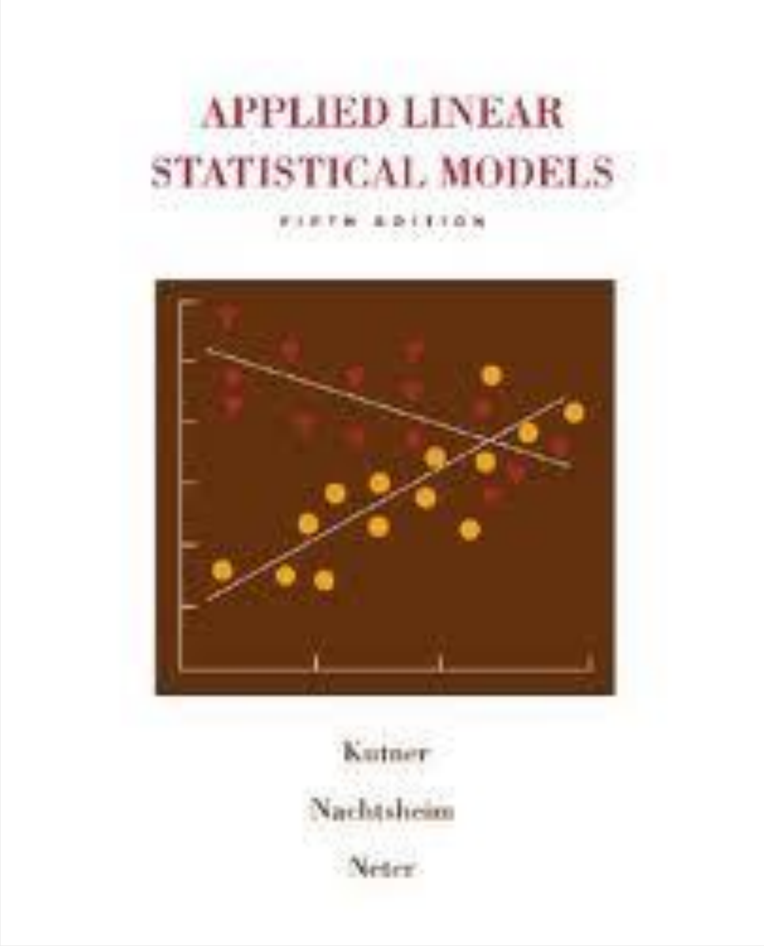
Paradigms for DOE

Classical:

- Fisher paradigm
- Textbook paradigm
- duPont paradigm

Optimal:

- Problem-driven paradigm



The Classical Approach to DOE

1. State the objectives and your constraints
2. Find a design from a book or a table or a software program that most closely matches your objectives (requires DOE expertise)
3. Implement the design
4. Analyze results

Classical design paradigms pros, cons

Pros

1. Designs are tried & true
2. Easy analysis

Classical design paradigms pros, cons

Pros

1. Designs are tried & true
2. Easy analysis

Cons

1. Requires much expertise
2. Rigidity: designs often do not match the problem

What DOE Expertise is Needed?

- Full factorial designs

What DOE Expertise is Needed?

- Full factorial designs
- Interactions

What DOE Expertise is Needed?

- Full factorial designs
- Interactions
- 2^k factorial designs

What DOE Expertise is Needed?

- Full factorial designs
- Interactions
- 2^k factorial designs
- 2^{k-p} fractional factorial designs

What DOE Expertise is Needed?

- Full factorial designs
- Interactions
- 2^k factorial designs
- 2^{k-p} fractional factorial designs
- Confounding schemes and aliasing

What DOE Expertise is Needed?

- Full factorial designs
- Interactions
- 2^k factorial designs
- 2^{k-p} fractional factorial designs
- Confounding schemes and aliasing
- Resolution

What DOE Expertise is Needed?

- Full factorial designs
- Interactions
- 2^k factorial designs
- 2^{k-p} fractional factorial designs
- Confounding schemes and aliasing
- Resolution
- Screening designs (Plackett Burman, Res III FF)

What DOE Expertise is Needed?

- Full factorial designs
- Interactions
- 2^k factorial designs
- 2^{k-p} fractional factorial designs
- Confounding schemes and aliasing
- Resolution
- Screening designs (Plackett Burman, Res III FF)

OK, What else?

What DOE Expertise (continued)?

- Response surface experiments

CARLSON
SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments
- Nested designs

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments
- Nested designs
- Repeated measures and split plot designs

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments
- Nested designs
- Repeated measures and split plot designs
- Incomplete block designs

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments
- Nested designs
- Repeated measures and split plot designs
- Incomplete block designs
- Mixture experiments and mixture models

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments
- Nested designs
- Repeated measures and split plot designs
- Incomplete block designs
- Mixture experiments and mixture models
- Supersaturated designs

What DOE Expertise (continued)?

- Response surface experiments
- Blocked experiments
- Nested designs
- Repeated measures and split plot designs
- Incomplete block designs
- Mixture experiments and mixture models
- Supersaturated designs

When might classical designs fail to match problem?

- **Multi-level designs: some factors at 2 levels, some at 3, some at four**

When might classical designs fail to match problem?

- **Multi-level designs: some factors at 2 levels, some at 3, some at four**
- **Constrained designs: Can’t run the factors simultaneously at the high levels, chemical process will blow up the plant!**

When might classical designs fail to match problem?

- **Multi-level designs: some factors at 2 levels, some at 3, some at four**
- **Constrained designs: Can’t run the factors simultaneously at the high levels, chemical process will blow up the plant!**
- **Some factors are qualitative, some not**

When might classical designs fail to match problem?

- **Multi-level designs: some factors at 2 levels, some at 3, some at four**
- **Constrained designs: Can't run the factors simultaneously at the high levels, chemical process will blow up the plant!**
- **Some factors are qualitative, some not**
- **I have a mixture experiment---factor levels are percentages that must add to one**

When might classical designs fail to match problem?

- **Multi-level designs: some factors at 2 levels, some at 3, some at four**
- **Constrained designs: Can't run the factors simultaneously at the high levels, chemical process will blow up the plant!**
- **Some factors are qualitative, some not**
- **I have a mixture experiment---factor levels are percentages that must add to one**
- **My block sizes cannot be the same.**

When might classical designs fail to match problem?

- Multi-level designs: some factors at 2 levels, some at 3, some at four
- Constrained designs: Can't run the factors simultaneously at the high levels, chemical process will blow up the plant!
- Some factors are qualitative, some not
- I have a mixture experiment---factor levels are percentages that must add to one
- My block sizes cannot be the same.
- I want to augment an existing design.

When might classical designs fail to match problem?

- Multi-level designs: some factors at 2 levels, some at 3, some at four
- Constrained designs: Can't run the factors simultaneously at the high levels, chemical process will blow up the plant!
- Some factors are qualitative, some not
- I have a mixture experiment---factor levels are percentages that must add to one
- My block sizes cannot be the same.
- I want to augment an existing design.
- I have a special model, etc., etc., etc.

New Paradigm: Problem-driven design

Problem-Driven Paradigm

- **With good software, approach easy to use and to teach.**
- **One, unified approach to (nearly) all design problems:**

1. Describe your responses
2. Describe your factors
3. Describe your objectives (model terms)
4. Describe your constraints and budget (*n*)
5. Press button to create the design
6. Check diagnostics, do sensitivity analysis
7. Manually alter as with classical design

CARLSON
SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

Problem-driven paradigm pros, cons:

Pros

1. Design tailored to problem
2. One unified approach to nearly all problems

Problem-driven paradigm pros, cons:

Pros

1. Design tailored to problem
2. One unified approach to nearly all problems

Cons

1. No textbook advocating computer-aided design
2. Technical concerns from the stats community

Problem-driven paradigm pros, cons:

Pros

1. Design tailored to problem
2. One unified approach to nearly all problems

Cons

1. No textbook advocating computer-aided design
2. Technical concerns from the stats community

Fading fast

Six Illustrations with JMP

Example 1: Extracting food solids from peanuts in solution, n = 16

	Factor	Low	High
1	Water pH level	6.95	8.0
2	Water temp	20C	60C
3	Extraction time	15	40
4	Water-Peanuts Ratio	5	9
5	Agitation speed	5,000	10,000
6	Hydrolyzed?	N	Y
7	Presoaking?	N	Y

Classical Approach: 2^{7-3} Resolution IV FF

Problem:

- Seven two-level factors in 16 runs.
- Main effects are critical
- Two-factor interactions might be present, might not.
- Standard approach: Resolution IV fractional factorial design

Well OK, but how did I happen to know this?

Knowledge Requirements:

- Main effects, interactions (model terms)
- Confounding schemes, aliasing
- Resolution
- Fractional factorials
- Plackett-Burman designs
- Blocking (maybe)

Contrast: Custom Design Approach

Knowledge Requirements:

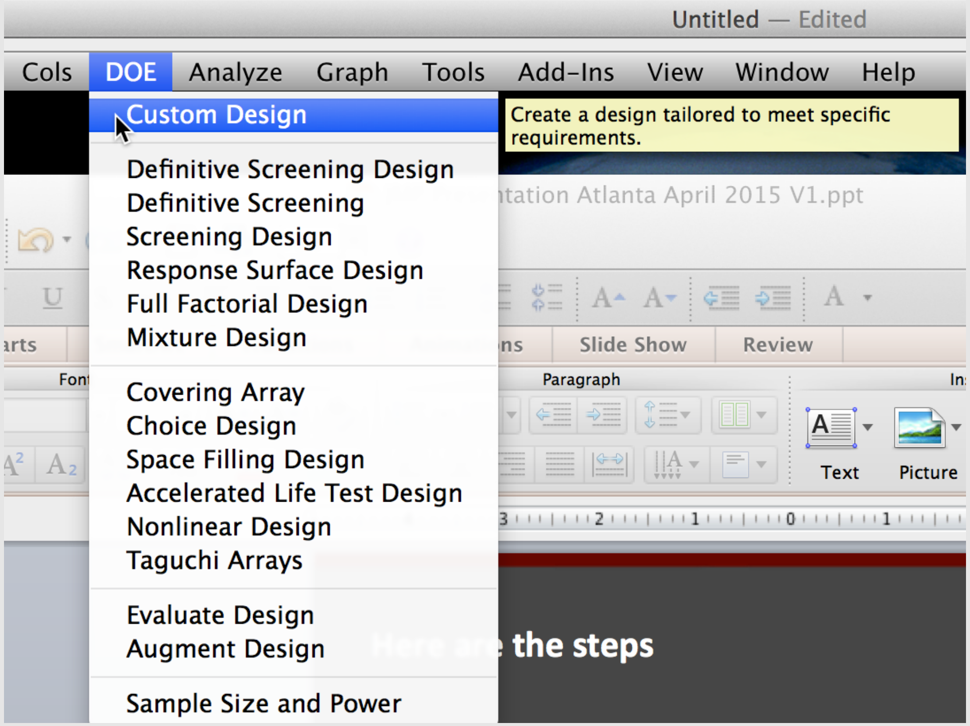
1. Main effects, interactions (model terms)
2. Knowledge of where the “button” is.

Here's how it's done:

- Specify the main effects model
- Give the sample size as usual
- Push the button

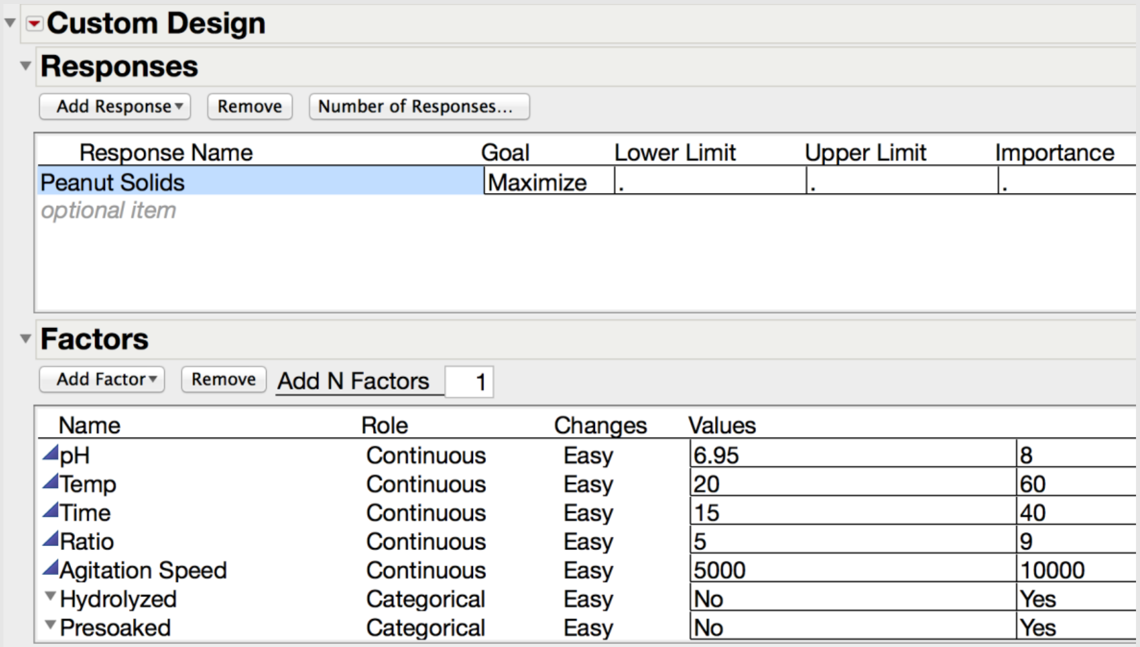
Here are the steps (Step 1)

1. DOE >> Custom Design



Here are the steps (Steps 2 and 3)

1. DOE >> Custom Design
2. Give the response
3. Give the factors levels (or ranges)



Custom Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance
Peanut Solids	Maximize	.	.	.
optional item				

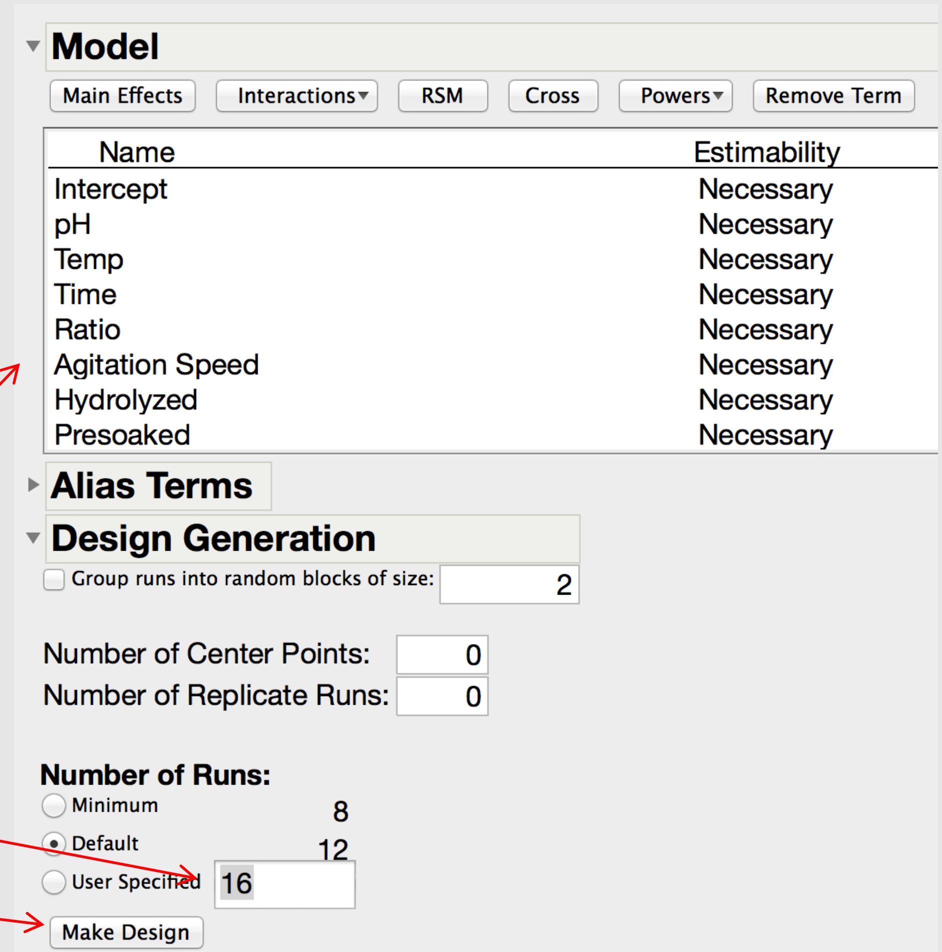
Factors

Add Factor Remove Add N Factors 1

Name	Role	Changes	Values	
pH	Continuous	Easy	6.95	8
Temp	Continuous	Easy	20	60
Time	Continuous	Easy	15	40
Ratio	Continuous	Easy	5	9
Agitation Speed	Continuous	Easy	5000	10000
Hydrolyzed	Categorical	Easy	No	Yes
Presoaked	Categorical	Easy	No	Yes

Here are the steps (Steps 4, 5 and 6)

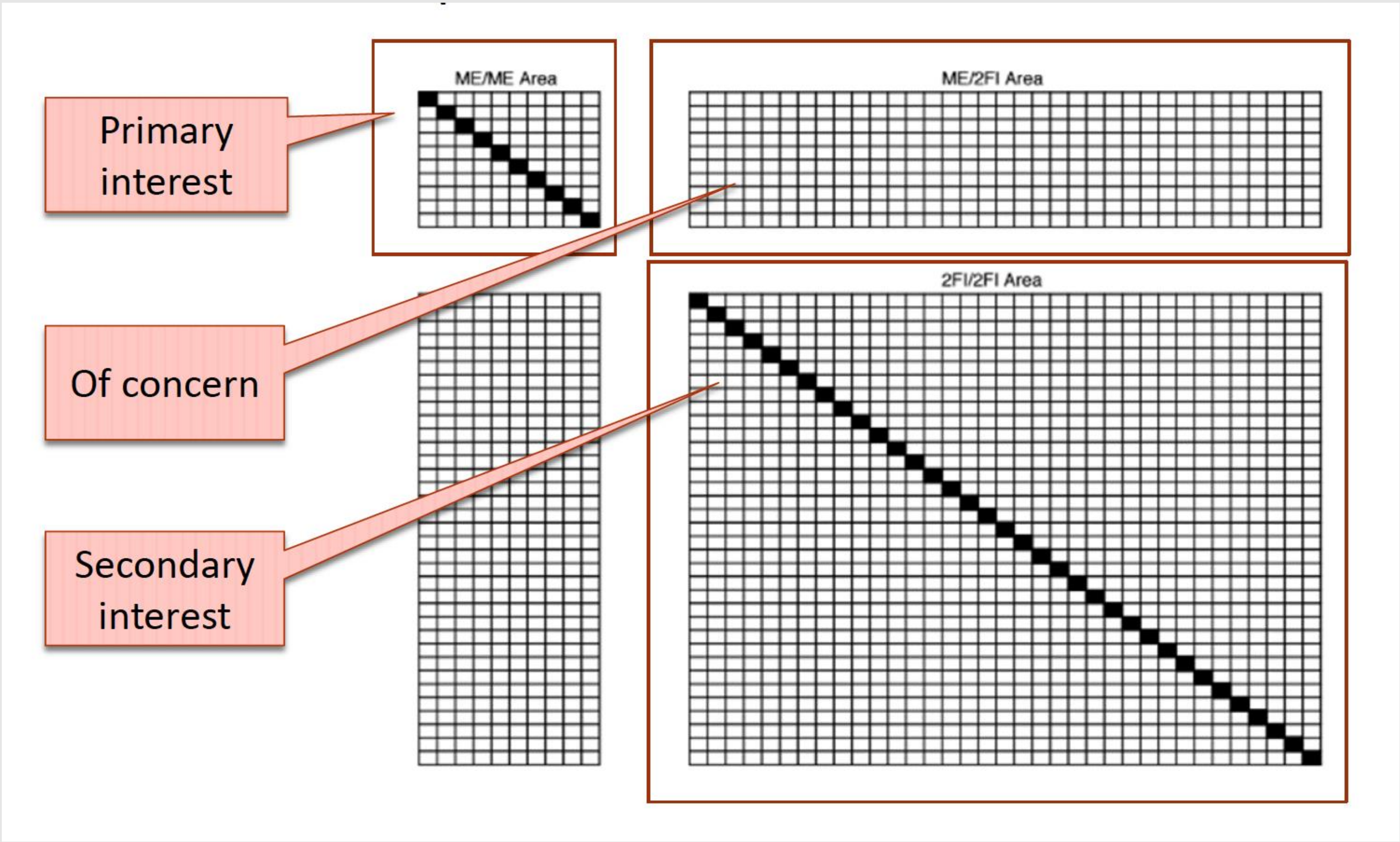
1. DOE >> Custom Design
2. Give the response
3. Give the factors and levels
4. JMP specifies the model
5. Give sample size
6. Press the button



The result: Orthogonal, D-optimal 16-Run Design

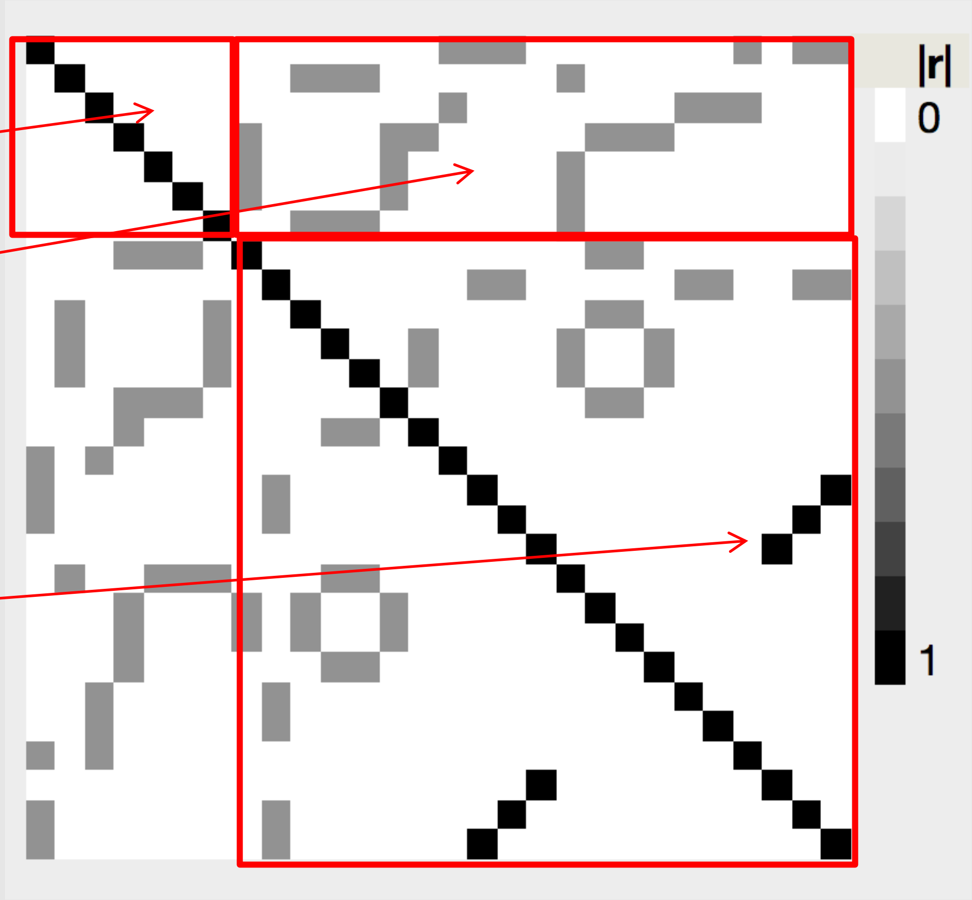
	pH	Temp	Time	Ratio	Agitation Speed	Hydrolyzed	Presoaked	Peanut Solids
1	6.95	60	15	5	5000	Yes	No	•
2	6.95	60	40	5	10000	No	No	•
3	8	20	15	5	5000	Yes	Yes	•
4	6.95	20	40	5	10000	Yes	No	•
5	6.95	60	40	9	5000	No	Yes	•
6	6.95	20	40	9	5000	Yes	Yes	•
7	6.95	20	15	9	10000	No	Yes	•
8	8	60	40	9	5000	Yes	No	•
9	8	60	15	5	10000	Yes	Yes	•
10	8	20	15	9	5000	No	No	•
11	8	20	40	5	10000	No	Yes	•
12	8	60	15	9	10000	No	No	•
13	8	60	40	9	10000	Yes	Yes	•
14	8	20	40	5	5000	No	No	•
15	6.95	20	15	9	10000	Yes	No	•
16	6.95	60	15	5	5000	No	Yes	•

One way to evaluate the design: Correlation Cell Plot



Our Design

1. Primary interest **(Yay!)**
2. Of concern **(Some correlations)**
1. Secondary interest **(Some two-factor interactions directly confounded, Boo!)**



Can we ask more of our design?

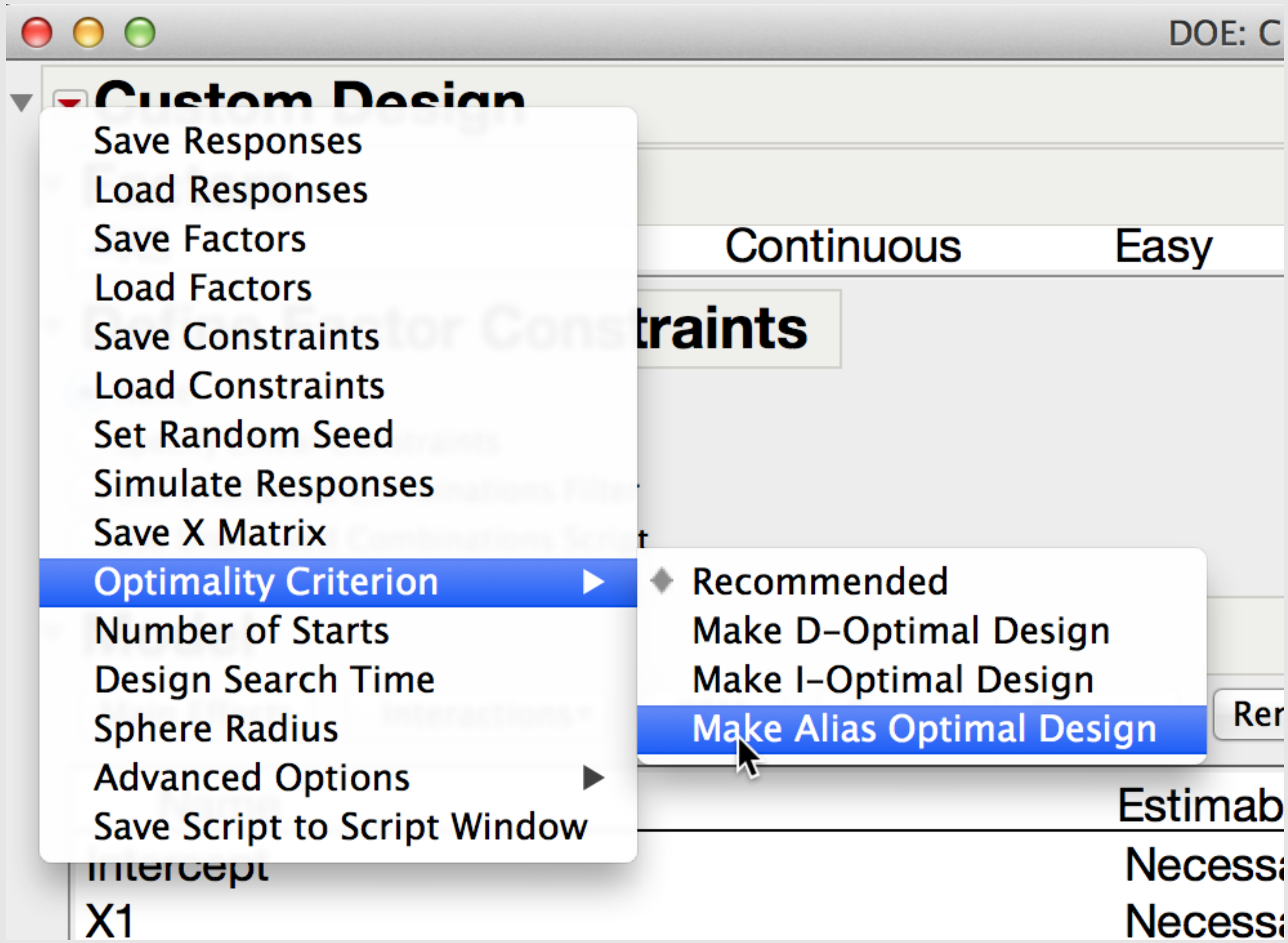
- Let’s tell JMP to find a design that minimizes the aliasing (correlations) between main effects and two-factor interactions

Only need to change the optimality criterion, and re-do

Red triangle >>

Optimality Criterion >>

Make Alias Optimal Design

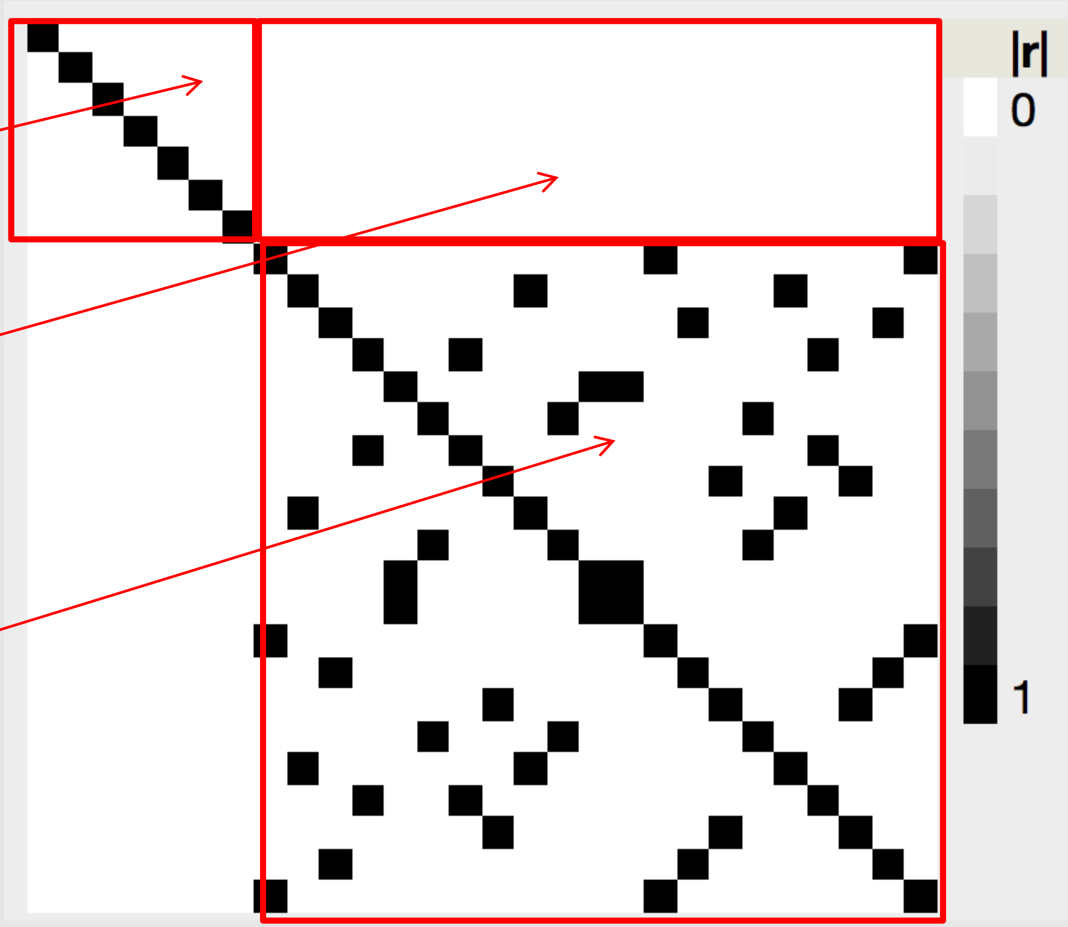


Result: Resolution IV fractional factorial design

1. Primary interest
   (Yay!)

2. Of concern
   (Yay!)

1. Secondary interest
   (Some two-factor interactions directly confounded)



CARLSON
SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

How did JMP do that?

Efficient Designs With Minimal Aliasing

Bradley JONES

SAS Institute
Cary, NC 27513
and
Faculty of Applied Economics
Universiteit Antwerpen
Prinsstraat 13
2000 Antwerpen, Belgium
(*Bradley.Jones@jmp.com*)

Christopher J. NACHTSHEIM

Operations and Management Science Department
Carlson School of Management
University of Minnesota
Minneapolis, MN 55455
(*Nacht001@umn.edu*)

For some experimenters, a disadvantage of the standard optimal design approach is that it does not consider explicitly the aliasing of specified model terms with terms that are potentially important but are not included in the model. For example, when constructing an optimal design for a first-order model, aliasing of main effects and interactions is not considered. This can lead to designs that are optimal for estimation of the primary effects of interest, yet have undesirable aliasing structures. In this article, we construct exact designs that minimize the squared norm of the alias matrix subject to constraints on design efficiency. We demonstrate use of the method for the construction of screening and response surface designs.

Technometrics, 2011, Vol 53, pp 62--71

JMP Demo: Peanut Solids…Over to Brad

Contrasts

Term	Contrast	Lenth t-Ratio	Individual p-Value	Simultaneous p-Value	Aliases
Ratio	-1.95810	-11.36	<.0001*	0.0004*	Pre-soak?*pH*Water Temp
Pre-soak?	-0.76222	-4.42	0.0041*	0.0396*	
Agitation Speed	0.68295	3.96	0.0064*	0.0598	
Hydrolyze?	-0.41454	-2.41	0.0319*	0.3035	
pH	0.11962	0.69	0.4718	1.0000	Ratio*Pre-soak?*Water Temp
Water Temp	0.00295	0.02	0.9888	1.0000	Ratio*Pre-soak?*pH
Extraction Time	0.00113	0.01	0.9959	1.0000	
Ratio*Ratio	0.01944 *	0.11	0.9192	1.0000	
Ratio*Pre-soak?	0.33411	1.94	0.0697	0.5543	pH*Water Temp
Ratio*Agitation Speed	0.11019	0.64	0.5591	1.0000	
Pre-soak?*Agitation Speed	-0.17147	-0.99	0.3061	0.9963	
Ratio*Hydrolyze?	0.29993 *	1.74	0.0922	0.6848	Agitation Speed*pH
Pre-soak?*Hydrolyze?	-0.06187 *	-0.36	0.7398	1.0000	Agitation Speed*Water Temp
Agitation Speed*Hydrolyze?	-0.29757 *	-1.73	0.0942	0.6955	Ratio*pH, Pre-soak?*Water Temp
Pre-soak?*pH	0.01945	0.11	0.9192	1.0000	Ratio*Water Temp
Pre-soak?*Extraction Time	0.14833 *	0.86	0.3662	0.9997	
Ratio*Pre-soak?*Hydrolyze?	-0.10430 *	-0.61	0.5814	1.0000	Pre-soak?*Agitation Speed*pH, Ratio*Agitation Speed*Water Temp

Example 2: Maximize recall of TV commercial

Factors:

A: Length of commercial
(10 to 30 seconds)

B: Number of repetitions
(1 to 3)

C: Amount of time since last viewing (1 to 5 days)

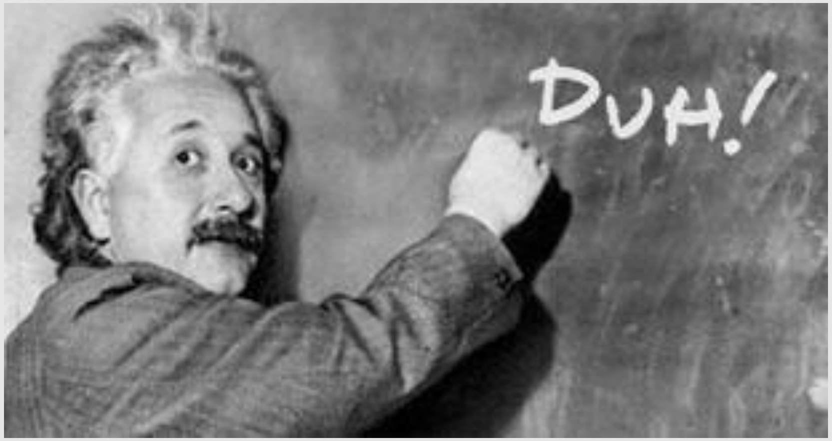
Classical Design Knowledge Requirements:

- Main effects, interactions, second-order powers (model terms)
- Confounding and resolution (for corners)
- Fractional factorials (for corners)
- CCDs, Box-Behnken designs
- Blocking (sort of)
- Axial points

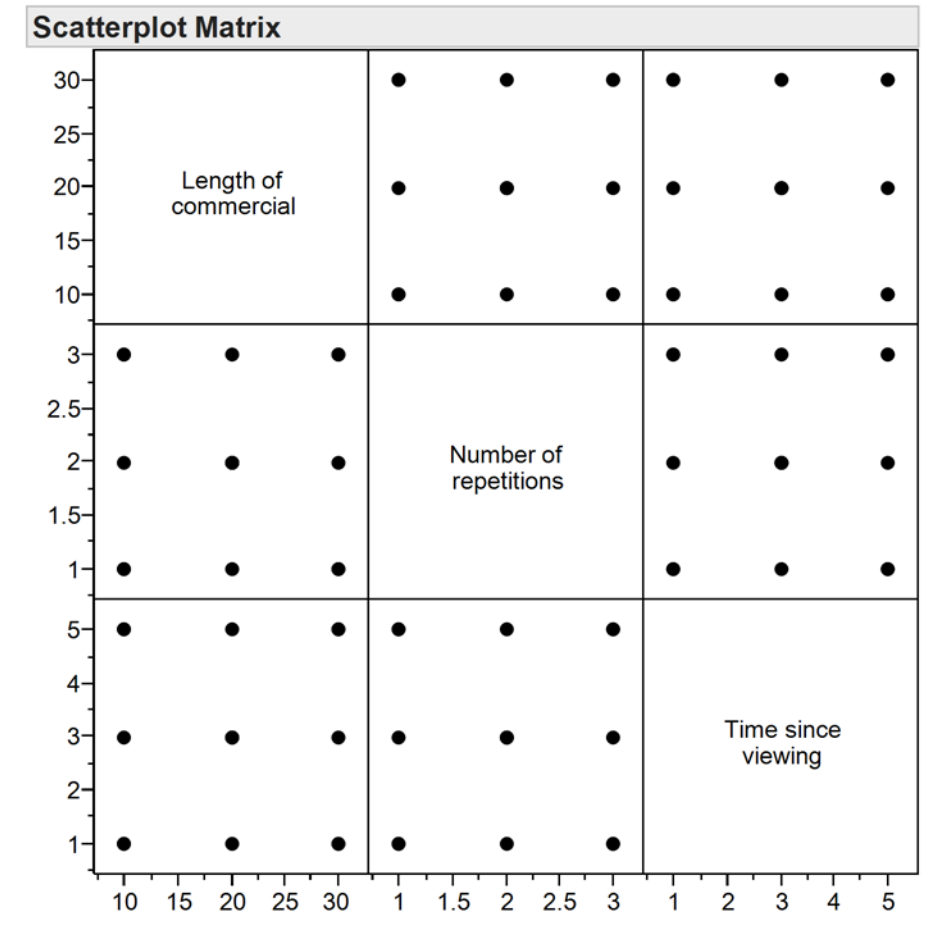
Contrast: Problem-Driven Design Approach

Knowledge Requirements:

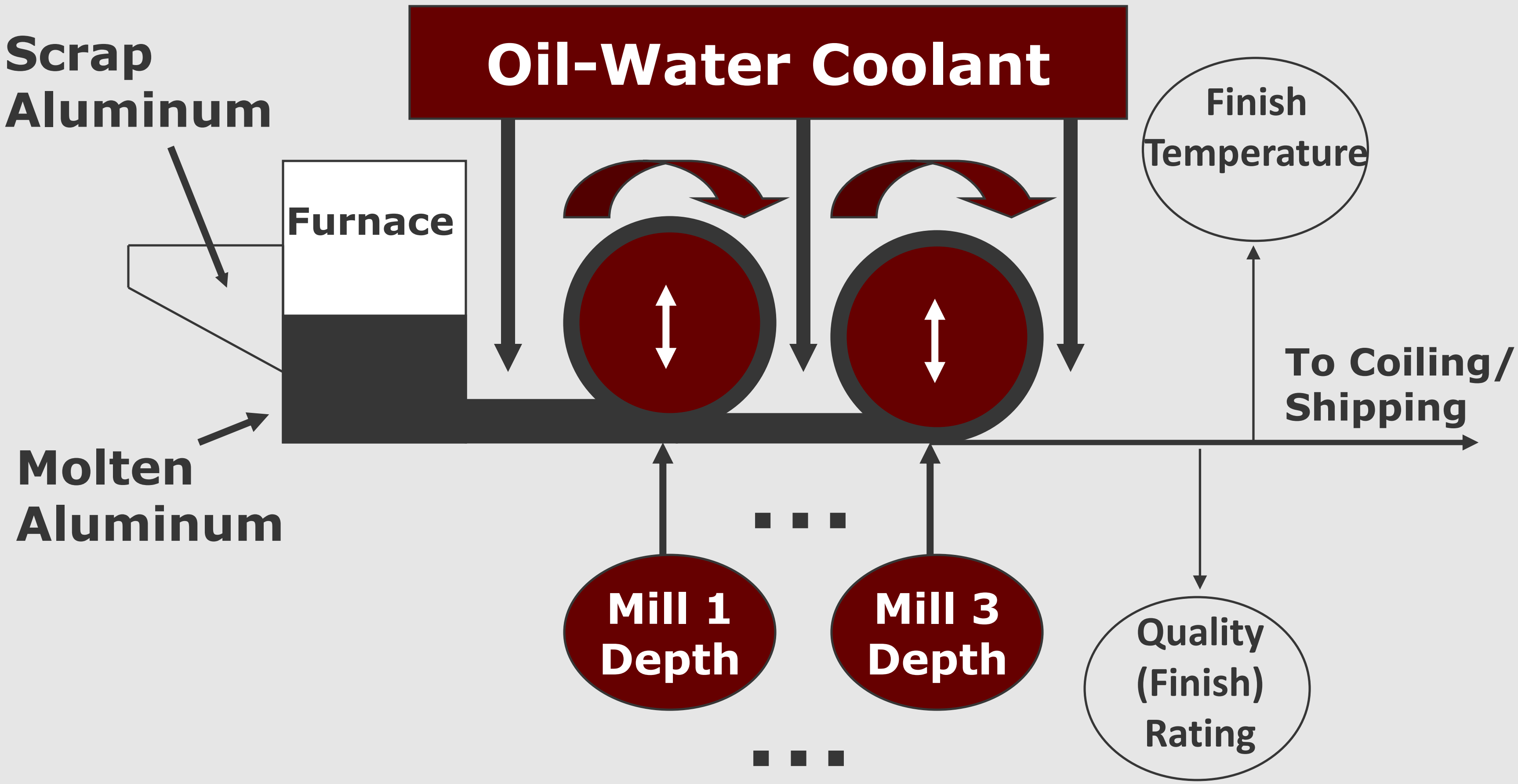
1. Main effects, interactions, quadratic terms (RSM button)



JMP Demo: Over to Brad



Example 3:
Sheet Aluminum Rolling Mill



Textbook approach to Aluminum Milling

	Factor	Low	High
1	Mill depth 1	0.1	0.8
2	Mill depth 2	0.1	0.8
3	Mill depth 3	0.1	0.8
4	Spray volume	L	H
5	Spray oil content	L	H

But there is a problem:

We have constrained mill depths (MDs):

$$MD1 + MD2 + MD3 = 1$$

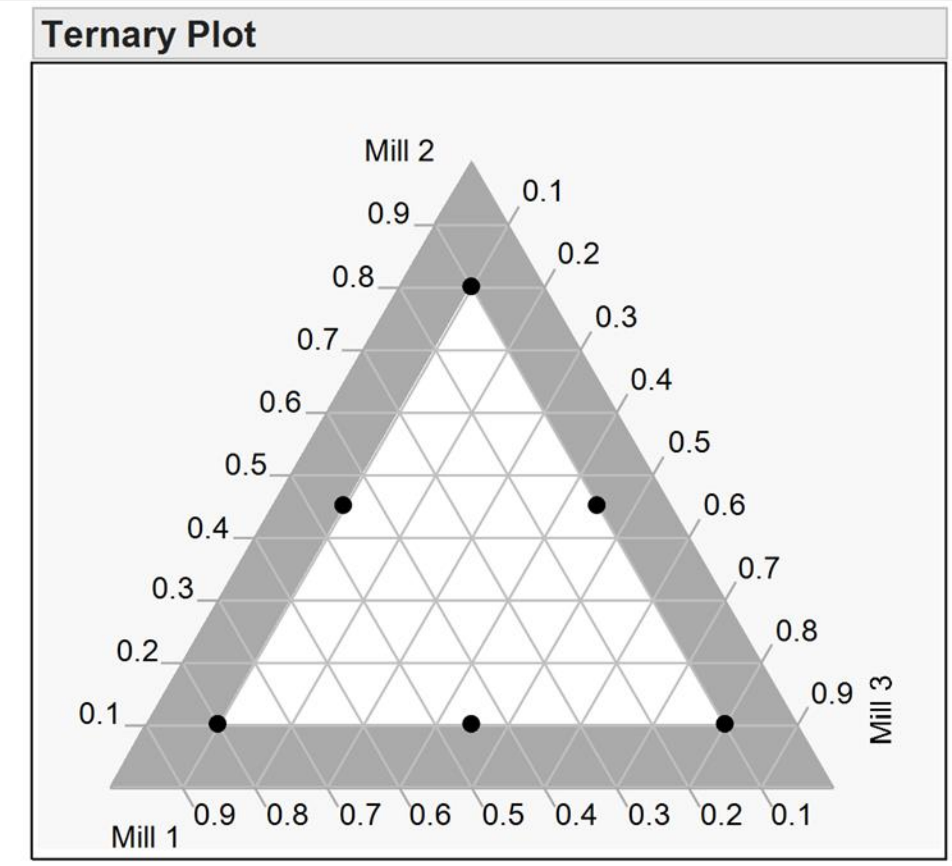
$$.1 < MD1 < .8$$

$$.1 < MD2 < .8$$

$$.1 < MD3 < .8$$

None of the tabled designs work!

JMP Demo: Over to Brad



DOE to the rescue!

- Result: When mill depth 3 increased, it improved the quality of the surface

Scrap reduced from 50% to 10%;

Company saved!

Split Plots: DOE with Easy and Hard to Change Factors

Very common in industrial experiments!

Can trip up even professional statisticians

"All industrial experiments are split-plot experiments"

--Cuthbert Daniel

Confession of a Text Book Author

- **I have only recently appreciated how effective split plot experiments really are.**
- **Thorough treatment should be standard in any introductory design course, any text, any statistics package.**

Confession of a Text Book Author

- I have only recently appreciated how effective split plot experiments really are.
- Thorough treatment should be standard in any introductory design course, any text, any statistics package.
- They merited a whopping three pages in my text book!

Recent Paper---My Conversion

Split-Plot Designs: What, Why, and How

BRADLEY JONES

SAS Institute, Cary, NC 27513

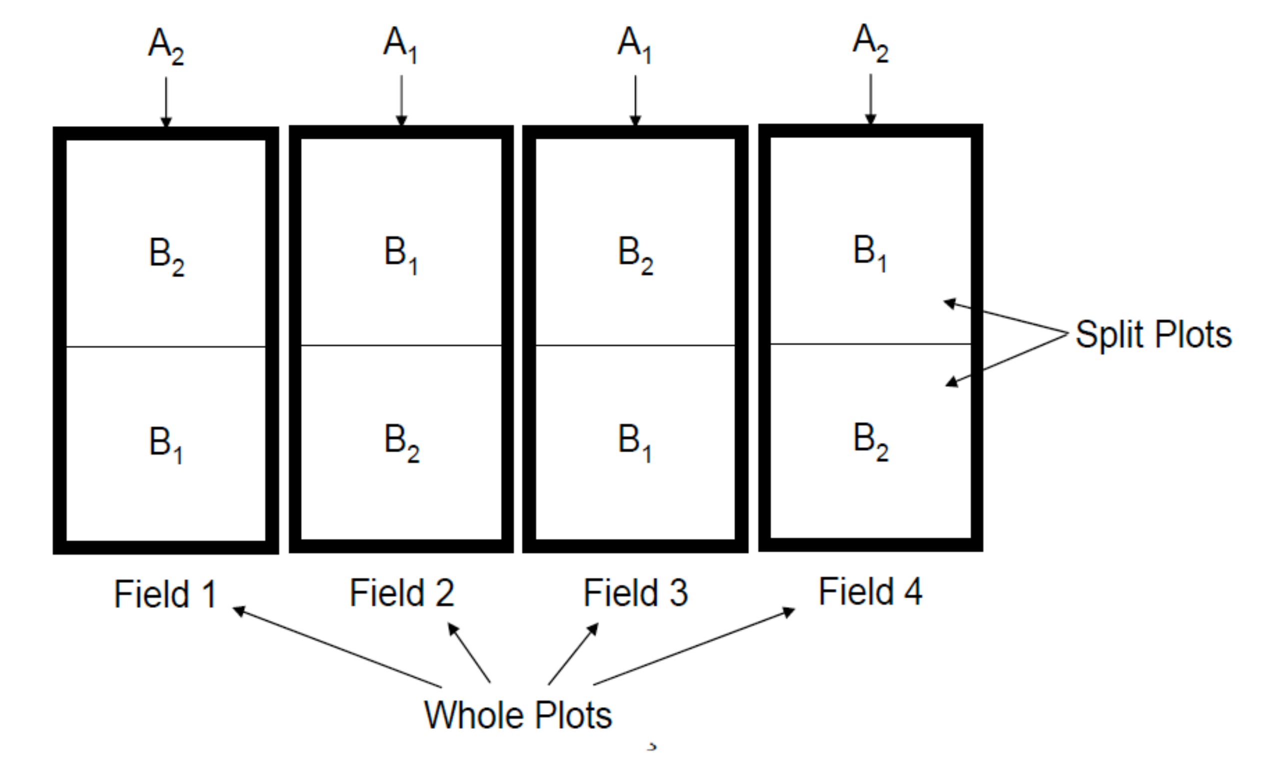
CHRISTOPHER J. NACHTSHEIM

Carlson School of Management, University of Minnesota, Minneapolis, MN 55455

Journal of Quality Technology 340 *Vol. 41, No. 4, October 2009*

So What is a Split-Plot Design and Why the Name?

Invented by Fisher for Agricultural Applications



Example 4: Corrosion resistance of steel bars (Box, Hunter, and Hunter text book)

DOE with two factors:

1. Furnace (curing) temperatures (360ºC, 370ºC, 380ºC)
2. Coating type (C_1, C_2, C_3, C_4)

Easy-to-change and hard-to-change factors

Problem: Not all factors are created equally!

Furnace temperature:

Hard to change!

Coating type:

Easy to change!

Actual Corrosion Resistance DOE

Whole-Plots (H2C)		Split-Plots (E2C)			
Furnace Run	Temperature	Coating in randomized order			
1	360	C2	C1	C3	C4
2	380	C1	C3	C4	C2
3	370	C3	C1	C2	C4
4	380	C4	C3	C2	C1
5	370	C4	C1	C3	C2
6	360	C1	C4	C2	C3

Advantage: *n* = 24 but only 6 furnace runs!

Key Split Plot DOE Characteristic

Two randomizations:

Step 1: Randomly assign temps to furnace runs

Step 2: For each furnace run…

Randomly assign coatings to four subgroups of bars within the furnace run

Contrast with ordinary DOE

In usual DOE: just one randomization:

3x4 factorial with an added replicate requires 24 furnace runs

(Only one coating type per furnace run)

3x4 Factorial

3x4 Factorial
Design 3x4 Factorial
Screening
Model

Columns (4/0)
Pattern
Temperature
Coating
Furnace Reset?

Rows
All rows 24
Selected 0
Excluded 0
Hidden 0
Labelled 0

	Pattern	Temperature	Coating
1	22	370	C2
2	13	360	C3
3	24	370	C4
4	31	380	C1
5	11	360	C1
6	32	380	C2
7	34	380	C4
8	34	380	C4
9	14	360	C4
10	31	380	C1
11	24	370	C4
12	23	370	C3
13	33	380	C3
14	13	360	C3
15	21	370	C1
16	21	370	C1
17	22	370	C2
18	23	370	C3
19	33	380	C3
20	12	360	C2
21	12	360	C2
22	32	380	C2
23	11	360	C1
24	14	360	C4

All-Too-Common Mistake

Step 1. Create ordinary randomized DOE

Step 2. Sort the runs by the hard-to-change factor and run in that order
(DOE is actually now a split plot design)

Step 3. Analyze as though a randomized DOE

Analysis depends on the randomization

Usual analysis	Factor	p-value	Significant?	
	Temperature	0.003	Yes	Wrong!
	Coatings	0.386	No	Wrong!
	Temp x Coatings	0.852	No	Wrong!

Analysis depends on the randomization

Usual analysis	Factor	p-value	Significant?	
	Temperature	0.003	Yes	Wrong!
	Coatings	0.386	No	Wrong!
	Temp x Coatings	0.852	No	Wrong!

Correct analysis	Factor	p-value	Significant?	
	Temperature	0.209	No	Right!
	Coatings	0.002	Yes	Right!
	Temp x Coatings	0.024	Yes	Right!

JMP Demo: Over to Brad

REML ANOVA Table

Source	Nparm	DF	DFDen	F Ratio	Prob > F
Temperature	2	2	3	2.7548	0.2093
Coating	3	3	9	11.4798	0.0020*
Temperature*Coating	6	6	9	4.3757	0.0241*

ANOVA Assuming Completely Randomized Design

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
Temperature	2	2	26519.250	10.2256	0.0026*
Coating	3	3	4289.125	1.1026	0.3860
Temperature*Coating	6	6	3269.750	0.4203	0.8518

Take-Aways

1. Split plots are easier to run!

1. Lose info on whole plot main effects, but…

1. Gain info on split-plot factors and split-plot by whole-plot factor interactions

2. Overall efficiency can be better than CRD

1. Teach these methods in workshops on DOE

What’s driving the software?

What’s driving the software?

Optimal design construction, optimal blocking

- Coordinate exchange algorithm, Meyer and Nachtsheim, *Technometrics*, 1995
- Optimal blocking, Cook, Nachtsheim, *Technometrics*, ‘1987

Bayesian D-Optimal Designs

- DuMouchel and Jones, *Technometrics*, 1994
- Jones, Lin, Nachtsheim,. *Journal of Statistical Planning and Inference*, 2006

Optimal Split-Plot and Split-Split Plot Designs

- Goos and Jones, *Biometrika*, 2009

Optimal designs depend on a model!

OH YEAH?

1. **No more so than classical designs.**

2. **They should!**

3. **If you're going to lose sleep, see model robustness, Bayesian, multi-objective design literature**

Lots of work in Model Robust and Multi-Objective Design

Randomly chosen examples

1. Li/Nachtsheim, 2000, Model Robust Factorial Designs

2. DuMouchel/Jones, 1994, Bayesian D-optimal Designs

3. Jones/Nachtsheim, 2011 Minimal Aliasing Designs

Definitive Screening Designs

CARLSON
SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

Multi-objective designs: six-factor example, n = 12

- **Standard choice: Placket-Burman design**

- **Plackett-Burman designs have “complex aliasing of the main effects by two-factor interactions.**

Alias Matrix for PB design

Full Model: 1 + 6 + 15 = 22 terms

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$$

Model we can fit:

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}^*$$

Leads to bias:

$$\mathrm{E}(\hat{\boldsymbol{\beta}}_1) = \boldsymbol{\beta}_1 + \mathbf{A}\boldsymbol{\beta}_2$$

Alias matrix:

$$\mathbf{A} = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2$$

Alias Matrix of Plackett-Burman design

Alias Matrix

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0.333	-0.33	0.333	0.333	0.333	-0.33	0.333	-0.33	0.333	-0.33
B	0	0.333	-0.33	0.333	0.333	0	0	0	0	-0.33	-0.33	-0.33	-0.33	0.333	0.333
C	0.333	0	0.333	-0.33	0.333	0	-0.33	-0.33	-0.33	0	0	0	-0.33	0.333	-0.33
D	-0.33	0.333	0	-0.33	0.333	-0.33	0	-0.33	0.333	0	-0.33	0.333	0	0	-0.33
E	0.333	-0.33	-0.33	0	-0.33	-0.33	-0.33	0	0.333	-0.33	0	-0.33	0	-0.33	0
F	0.333	0.333	0.333	-0.33	0	-0.33	0.333	0.333	0	0.333	-0.33	0	-0.33	0	0

PB (non-regular) design has "complex aliasing"

If only there were another design with this alias matrix:

Alias Matrix

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Turns out there is:

Six foldover pairs

Run	A	B	C	D	E	F
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0
13	0	0	0	0	0	0

Definitive Screening Design for 6 factors

Center point in each row

Run	A	B	C	D	E	F
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0
13	0	0	0	0	0	0

Definitive Screening Design for 6 factors

Run	A	B	C	D	E	F
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0
13	0	0	0	0	0	0

One overall center point

How did we find this design?*

Design criterion: Minimize $\|\mathbf{A}\|^2$ subject to an efficiency constraint on the design:

$$\min_d \mathrm{Tr}[\mathbf{A}(d)'\mathbf{A}(d)], \text{ subject to } D_e(d) \geq l_D$$

where the D-efficiency of the design is:

$$D_e(d) = \left[\frac{|\mathbf{X}_1(d)'\mathbf{X}_1(d)|}{|\mathbf{X}_1(d_D)'\mathbf{X}_1(d_D)|}\right]^{1/p_1}$$

*Jones, Nachtsheim, *Technometrics*, 2011

CARLSON SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

How did we find this design?*

- Minimize the average magnitude of the correlations between main effects and two-factor interactions
- Subject to a constraint on the statistical efficiency of the design (e.g., efficiency > 90%)

*Jones, Nachtsheim, *Technometrics*, 2011

Now generalize this structure for m factors

Table 1: General design structure for m factors

Foldover Pair	Run (i)	Factor Levels				
		$x_{i,1}$	$x_{i,2}$	$x_{i,3}$	$\cdots$	$x_{i,m}$
1	1	0	± 1	± 1	$\cdots$	± 1
	2	0	∓ 1	∓ 1	$\cdots$	∓ 1
2	3	± 1	0	± 1	$\cdots$	± 1
	4	∓ 1	0	∓ 1	$\cdots$	∓ 1
3	5	± 1	± 1	0	$\cdots$	± 1
	6	∓ 1	∓ 1	0	$\cdots$	∓ 1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
m	$2m-1$	± 1	± 1	± 1	$\cdots$	0
	$2m$	∓ 1	∓ 1	∓ 1	$\cdots$	0
Centerpoint	$m+1$	0	0	0	$\cdots$	0

Can we find great designs for any m?

It turns out there is a "Conference Matrix" solution

An *mxm* square matrix C with 0 diagonal and +1 or -1 off diagonal elements such that:

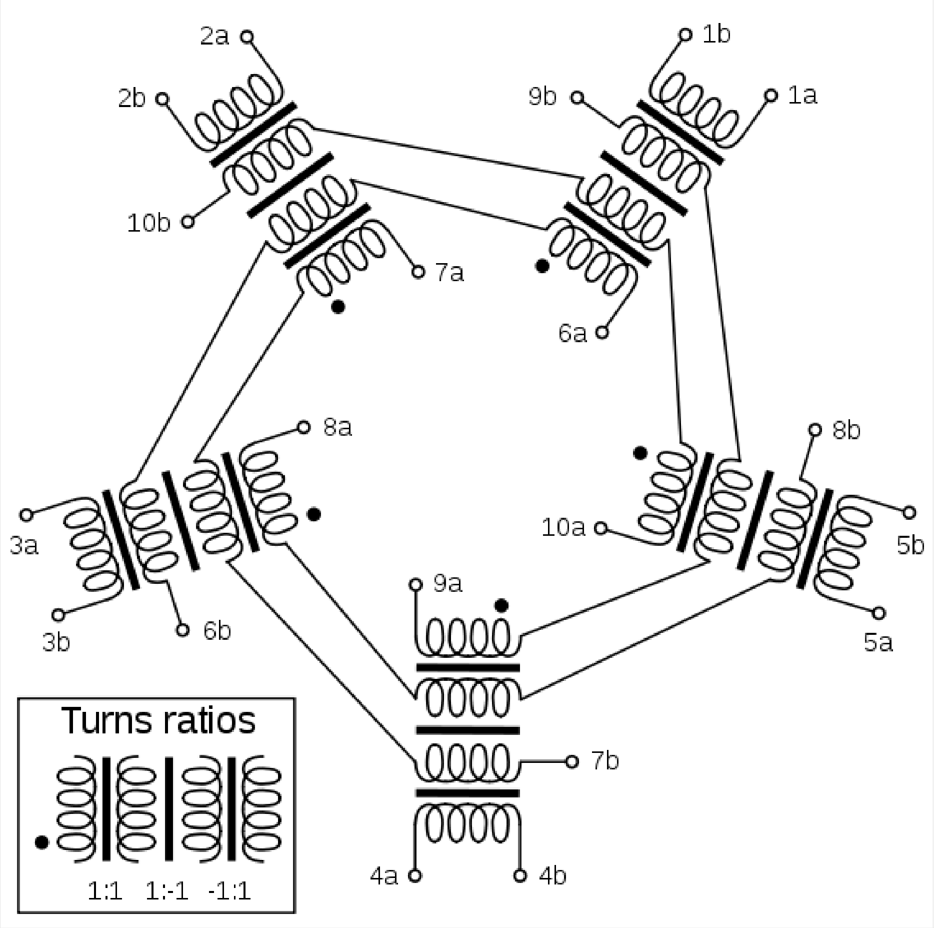
$$\mathbf{C}^{T}\mathbf{C} = (m - 1)\mathbf{I}_{m \times m}$$

- **Arose in connection with a problem in telephony and first described and named by Vitold Belevitch**
- **Belevitch was interested in constructing ideal telephone conference networks from ideal transformers and discovered that such networks were represented by C.**

Carlson School of Management University of Minnesota

Conference Matrices & Telephony

From Wikipedia article on Conference Matrices



Conference Matrix of Order 6

$$C = \begin{pmatrix} 0 & +1 & +1 & +1 & +1 & +1 \\ +1 & 0 & +1 & -1 & -1 & +1 \\ +1 & +1 & 0 & +1 & -1 & -1 \\ +1 & -1 & +1 & 0 & +1 & -1 \\ +1 & -1 & -1 & +1 & 0 & +1 \\ +1 & +1 & -1 & -1 & +1 & 0 \end{pmatrix}$$

Here the amazing result:

Form the augmented matrix:

$$D = \begin{bmatrix} +C \\ -C \\ 0 \end{bmatrix}$$

…and you get an orthogonal (for main effects) definitive screening design!

Design Properties

1. **The number of required runs is only one more than twice the number of factors.**

Design Properties

1. **The number of required runs is only one more than twice the number of factors.**
2. **Unlike Plackett-Burman and Resolution III and IV fractional factorial designs, main effects are completely independent of two-factor interactions.**

Design Properties

1. **The number of required runs is only one more than twice the number of factors.**
2. **Unlike Plackett-Burman and Resolution III and IV fractional factorial designs, main effects are completely independent of two-factor interactions.**
3. **Unlike all predecessors, these are three-level designs, so we can estimate curvatures!**

Design Properties

1. The number of required runs is only one more than twice the number of factors.
2. Unlike Plackett-Burman and Resolution III and IV fractional factorial designs, main effects are completely independent of two-factor interactions.
3. Unlike all predecessors, these are three-level designs, so we can estimate curvatures!
4. Designs are capable of estimating all possible response surface models involving three or fewer factors with very high levels of statistical efficiency.

Impact? A break-through solution for sequestering greenhouse gasses

2012 Statistics in Chemistry Award Winner

Bradley Jones, SAS Institute, JMP Division

Scott Allen, Novomer, Inc.

To Bradley Jones and Scott Allen for outstanding collaborative work in developing a new catalyst for CO_2-based polymers that sequester CO_2.

Impact? First published DSD case study, 2013

Biotechnol Lett
DOI 10.1007/s10529-012-1089-y

ORIGINAL RESEARCH PAPER

Efficient biological process characterization by definitive-screening designs: the formaldehyde treatment of a therapeutic protein as a case study

Axel Erler · Nuria de Mas · Philip Ramsey · Grant Henderson

Impact? From the conclusions:

"*Definitive-screening designs were used to efficiently select a model describing the formulation of a protein under clinical development. The ability of the single definitive screening design to identify and model all the active effects obviated the need for further experimentation, reducing the total number of experimental runs required to 17 from the greater than or equal to 70 runs that would have been necessary using the traditional screening/RSM approach.*"

Awards for Definitive Screening Designs

"A New Class of Three-Level Screening Designs for Definitive Screening in the Presence of Second-Order Effects", *Journal of Quality Technology*, **Jan., 2011**

- **Brumbaugh Award, 2011**
- **Lloyd S. Nelson Award, 2012**
- **(Led to) Statistics in Chemistry Award, 2012**
- **Much work now in combinatorial aspects**

Upshot – Definitive Screening Designs

1. Our view: engineers, scientists prefer three levels.
2. Can estimate curvatures
3. Can disentangle interactions
4. Designs project to fully estimable response surface designs in two or three factors, so with effect sparsity, two experiments for the price of one.
5. We see little or no reason to continue the practice of using 2^{k-p} designs for four or more continuous factors!!

BUT!...a limitation

DSDs can only be used with continuous (quantitative) factors

BUT!...a limitation

DSDs can only be used with continuous (quantitative) factors

DANG!!!!

Not any more:

Definitive Screening Designs with Added Two-Level Categorical Factors

BRADLEY JONES

SAS Institute, Cary, NC

CHRISTOPHER J. NACHTSHEIM

Carlson School of Management, University of Minnesota, Minneapolis, MN

- *Journal of Quality Technology*, April 2013

Example 5: Extracting food solids from peanuts in solution, again, n = 18

		Factor	Low	High
5 continuous factors	1	Water pH level	6.95	8.0
	2	Water temp	20C	60C
	3	Extraction time	15	40
	4	Water-Peanuts Ratio	5	9
	5	Agitation speed	5,000	10,000
2 categorical factors	6	Hydrolyzed?	N	Y
	7	Presoaking?	N	Y

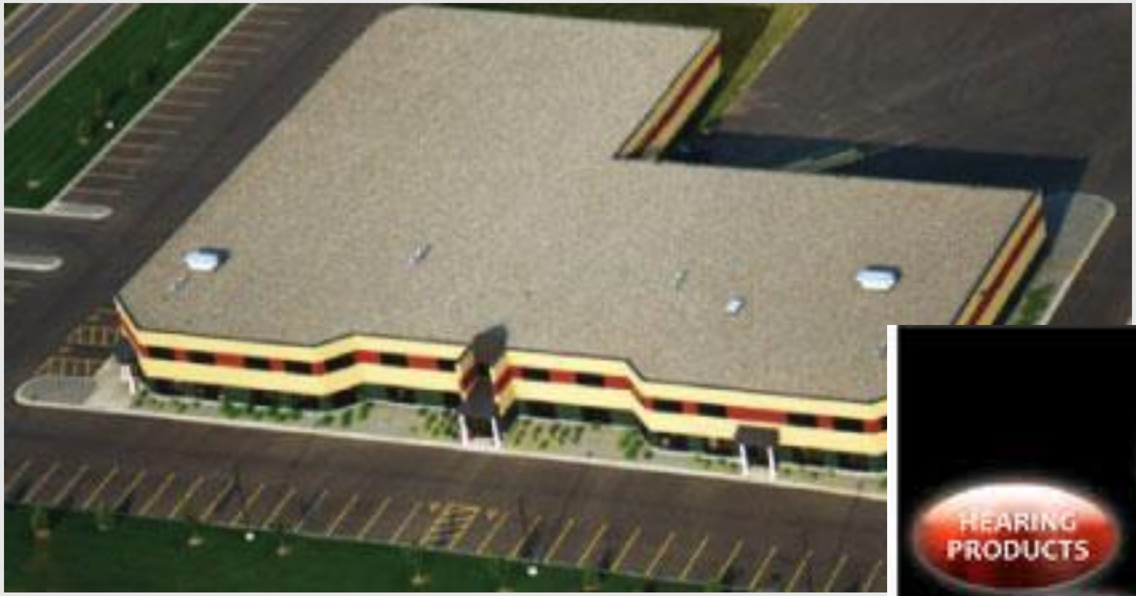
JMP Demo: Peanut Solids as a DSD---Over to Brad

Recall our conclusion: **There are three active interaction effects, but we cannot identify them!**

Using the same number of runs (16, plus 2 center points) the definitive screening design will not only identify the same main effects, but it will:

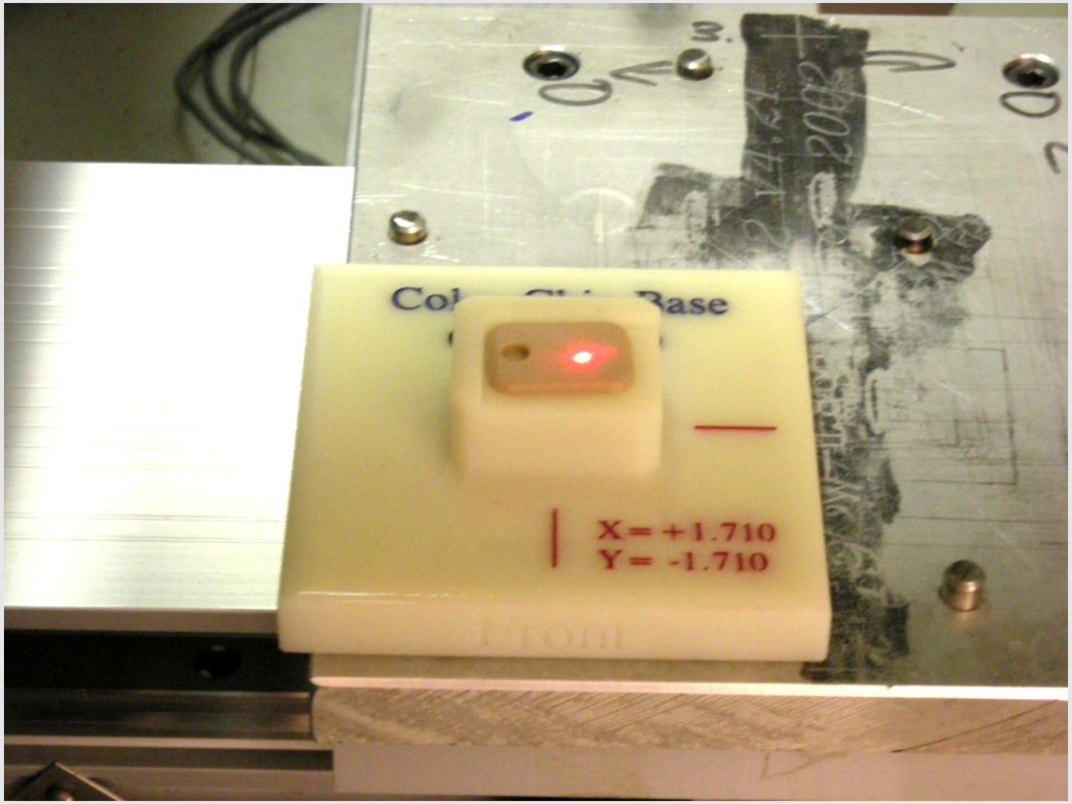
- Investigate all curvatures (concluding there are none)
- Investigate all interactions and **definitively** identify the three active interactions

Example 6: A Recent Experiment at In’Tech Industries



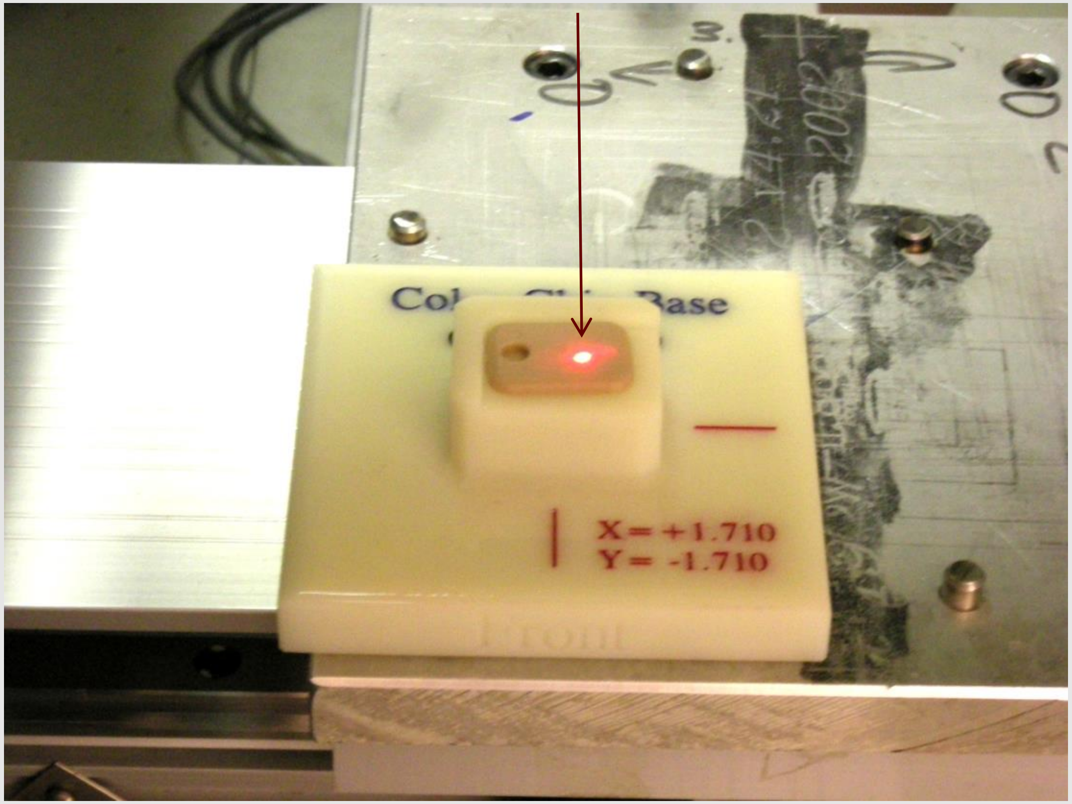
Problem: Need to laser etch labels on small plastic parts in an "optimal" fashion

Laser etching in progress….

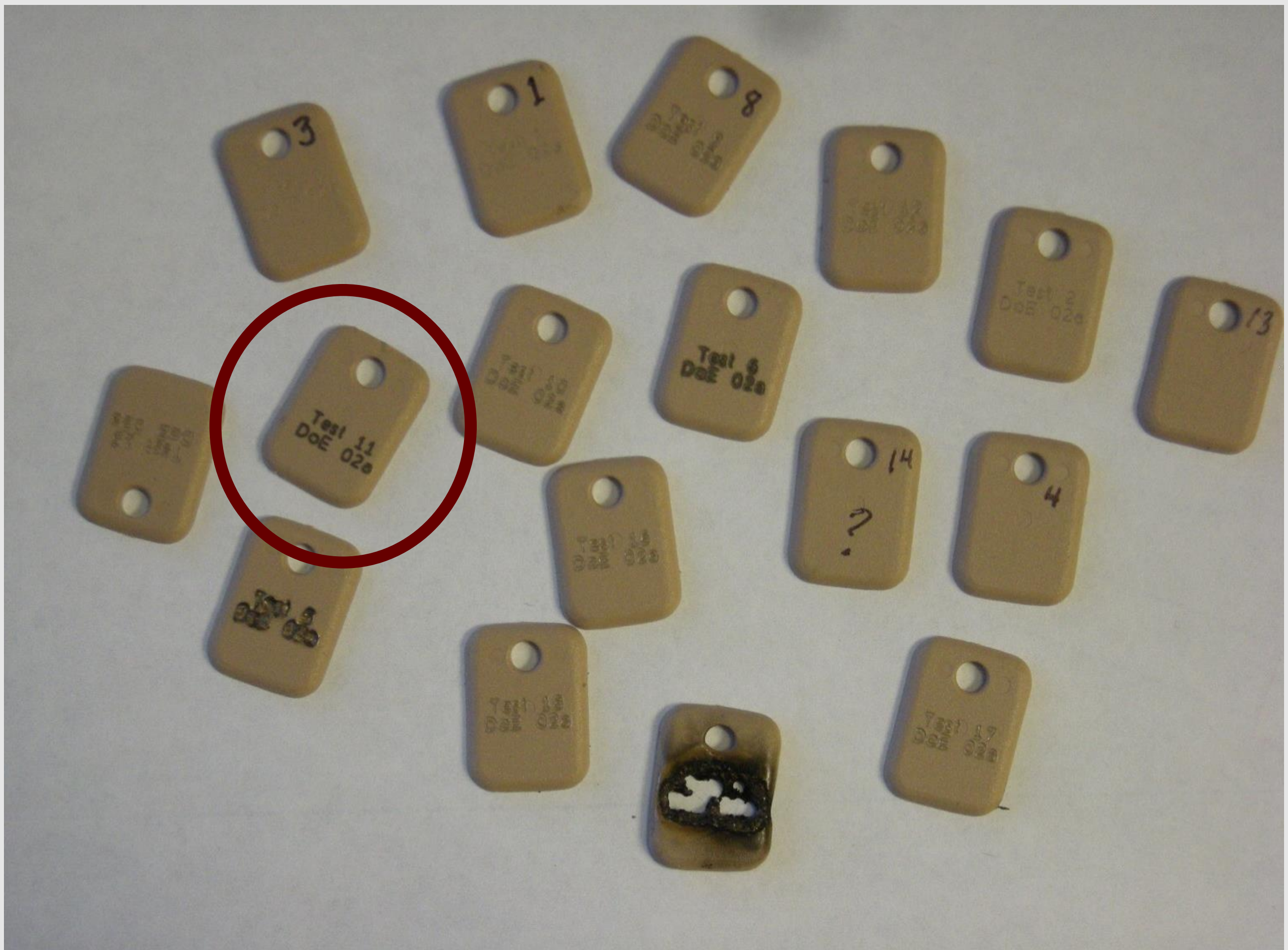


Laser etching in progress….

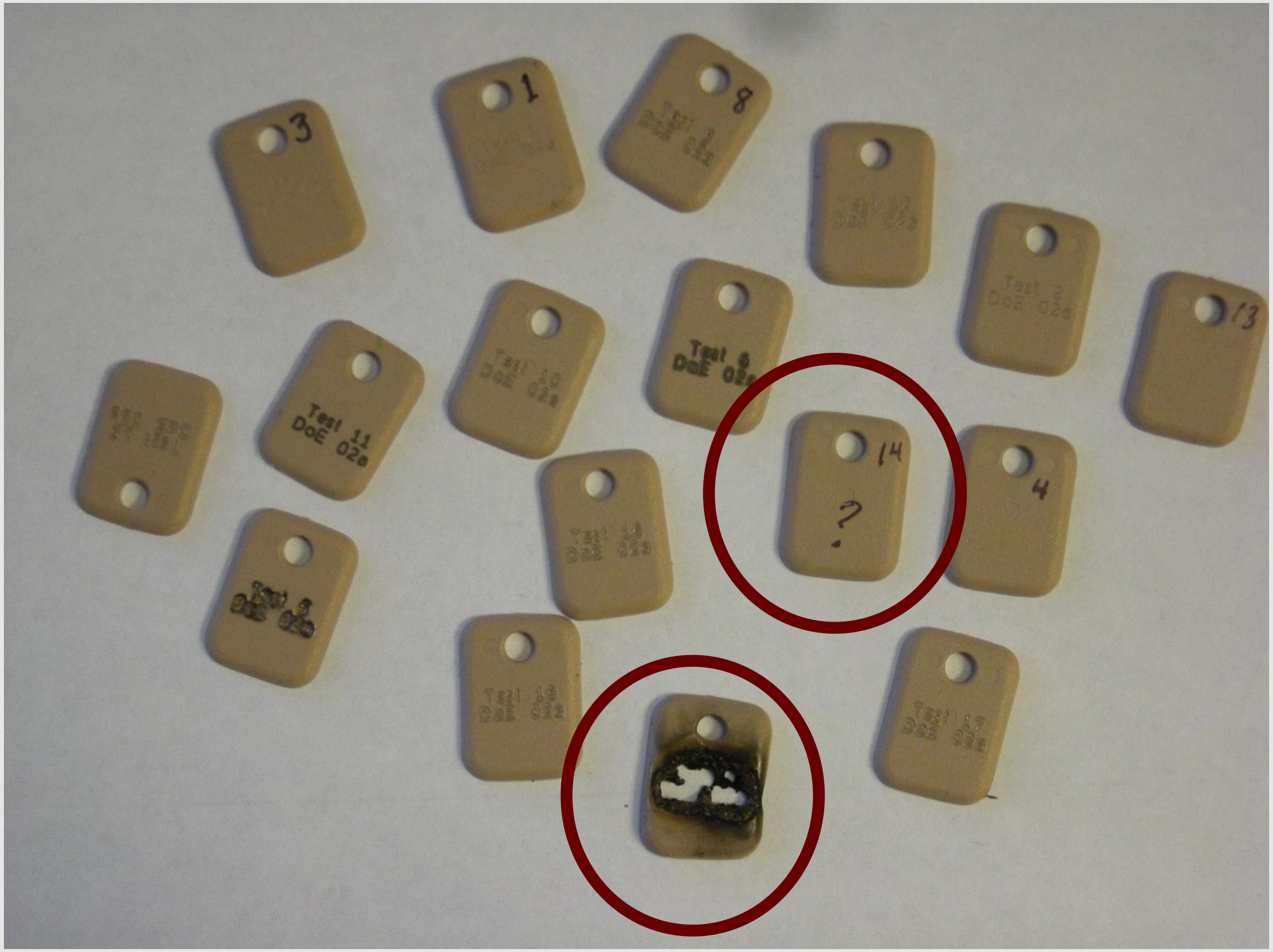
Laser



Initial experience: From easy to read,



Initial results: ……………… to not so easy to read



Factors and ranges....

Factors	Factor ranges	
	Low level	High level
Mark Speed	8	15
Frequency	1	5
Percent Power	15	55
Repetitions	1	5
Humidity	5%	15%

Blocking factor is operator

Ultimate objective: optimization

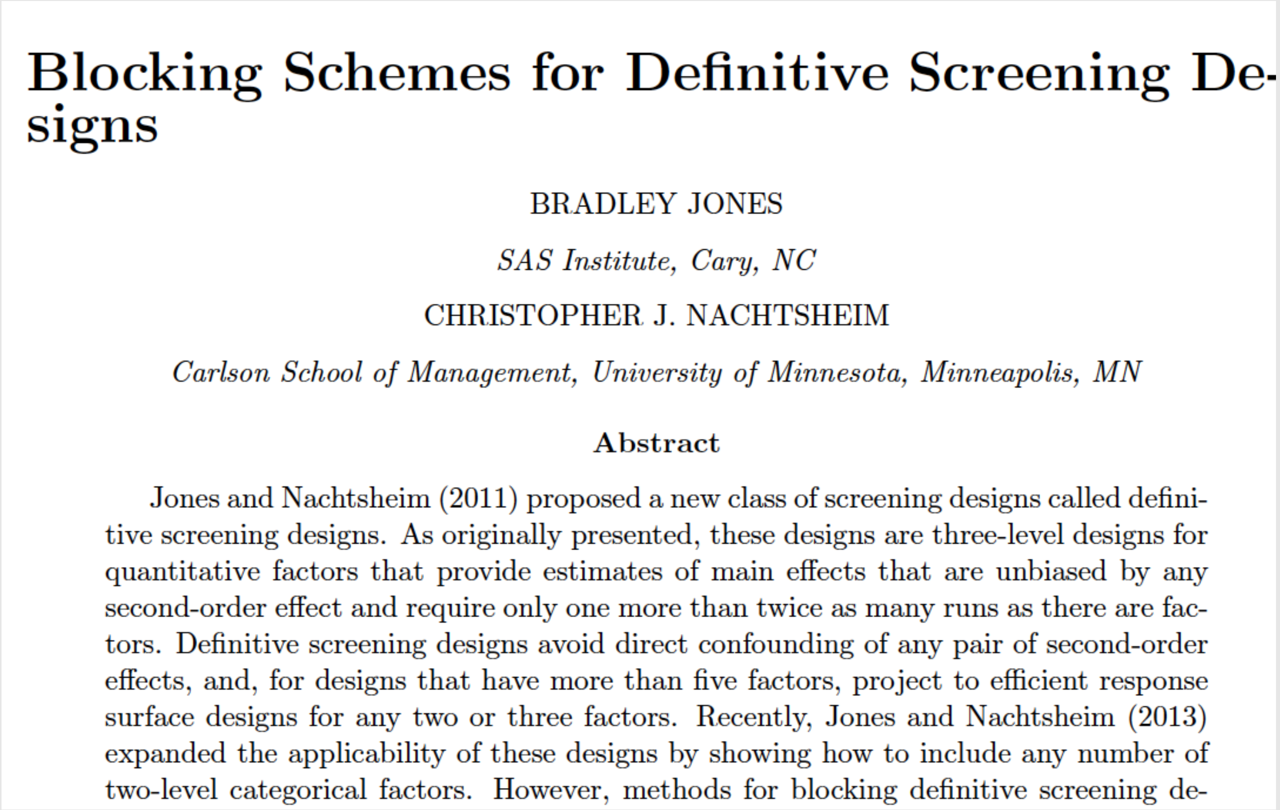
Experimental plan

Step 1: Screen (5 factors @ 3 levels - DSD)
Step 2: Optimize: RSD in fewer(?) factors

Complication: Need to block on operators

Conundrum: How to block a definitive screening design?

Solution to appear in Technometrics, 2015



Blocking Schemes for Definitive Screening Designs

BRADLEY JONES

SAS Institute, Cary, NC

CHRISTOPHER J. NACHTSHEIM

Carlson School of Management, University of Minnesota, Minneapolis, MN

Abstract

Jones and Nachtsheim (2011) proposed a new class of screening designs called definitive screening designs. As originally presented, these designs are three-level designs for quantitative factors that provide estimates of main effects that are unbiased by any second-order effect and require only one more than twice as many runs as there are factors. Definitive screening designs avoid direct confounding of any pair of second-order effects, and, for designs that have more than five factors, project to efficient response surface designs for any two or three factors. Recently, Jones and Nachtsheim (2013) expanded the applicability of these designs by showing how to include any number of two-level categorical factors. However, methods for blocking definitive screening de-

Analyzing DSDs

Simplest idea – Fit the main (linear) effects model

Advantages:

MEs unbiased – you can believe the coefficient estimates

Fit the Main Effects Model

Disadvantages:

Estimate of σ inflated with strong 2FIs or quadratic effects

May make active MEs appear not statistically significant.

You cannot believe the coefficient standard errors.

Analysis Idea #2 – Use Stepwise Regression

Advantages:

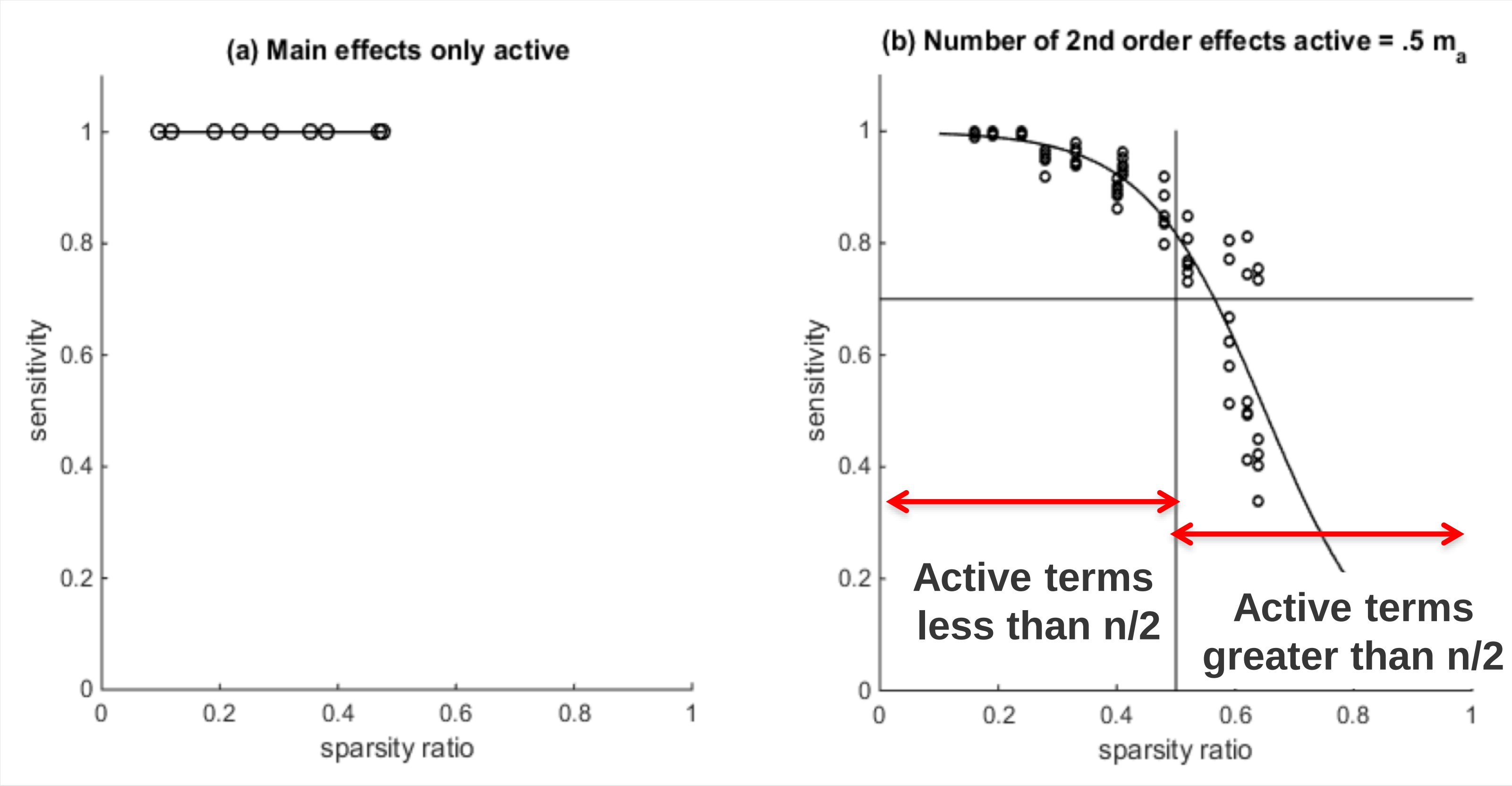
- Easy to do
- Available in most software
- Specify the response surface model
- Use forward variable selection based on the AICc criterion
- Or use the Dantzig Selector

Limitations of DSDs

1. Stepwise/AICc and Dantzig break down if there are more than about **n/2 active terms** in the model
2. The quadratic effect must be large (2-3 sigma) to have good power.

140

Simulation Results*



Fix 1: “Separate” Analyses of First-, Second-Order Effects

Since main effects and 2nd order effects are orthogonal to each other we can split the response (Y) into two new responses

- One response for identifying main effects – call it Y_ME
- One response for identifying the intercept and the 2nd order effects – call it Y_2nd
- And the two columns are orthogonal to each other

142

Fix 2: Employ two or more fake factors

- Adding fake factors provides a way to estimate variance without repeating center runs!
  - Fake factors are orthogonal to the real factors and all the 2nd order effects
  - Assuming the 3rd and higher order effects are negligible, we can use the fake factor degrees of freedom to create an unbiased estimate of the error variance!
- Use this estimate to identify active main effects

143

Analysis Start – Computing the New Responses

1. Fit the main effects model (no Intercept) to the real and fake factors and compute the residuals – these residuals are the responses for the 2nd order effects (Y2nd).
2. Save the predicted values from the above regression for the main effects response (YME).

144

Stage 1 – Identify Active Main Effects

1. Recall that the residuals from fitting the main effects data to the real factors only have 2 degrees of freedom.
2. To estimate σ^2 , sum the squared residuals from this fit and divide the result by 2.
3. Using this estimate, do t-tests of each coefficient and retain the active main effects

145

Stage 2 – Identify 2^{nd} Order Effects Assuming Heredity

- Form all the 2^{nd} order terms involving the active main effects
- Using Y_{2nd}, do forward all subsets regression up to the point where the MSE of the best 2^{nd} order model for a given number of terms is not significantly larger than your estimate of σ^2 from Stage 1

Example:

A	B	C	D	E	F	Fake 1	Fake 2	Y	Y_2nd	Y_ME
0	1	1	1	1	1	1	1	94.51	101.04	-6.53
0	-1	-1	-1	-1	-1	-1	-1	107.57	101.04	6.53
1	0	1	1	-1	1	-1	-1	94.36	101.1...	-6.815
-1	0	-1	-1	1	-1	1	1	107.99	101.1...	6.815
1	-1	0	1	1	-1	1	-1	91.80	90.525	1.275
-1	1	0	-1	-1	1	-1	1	89.25	90.525	-1.275
1	-1	-1	0	1	1	-1	1	93.70	94.485	-0.785
-1	1	1	0	-1	-1	1	-1	95.27	94.485	0.785
1	1	-1	-1	0	1	1	-1	89.55	88.71	0.84
-1	-1	1	1	0	-1	-1	1	87.87	88.71	-0.84
1	-1	1	-1	-1	0	1	1	94.58	95.235	-0.655
-1	1	-1	1	1	0	-1	-1	95.89	95.235	0.655
1	1	-1	1	-1	-1	0	1	93.23	89.58	3.65
-1	-1	1	-1	1	1	0	-1	85.93	89.58	-3.65
1	1	1	-1	1	-1	-1	0	98.11	95.815	2.295
-1	-1	-1	1	-1	1	1	0	93.52	95.815	-2.295
0	0	0	0	0	0	0	0	99.75	99.75	0

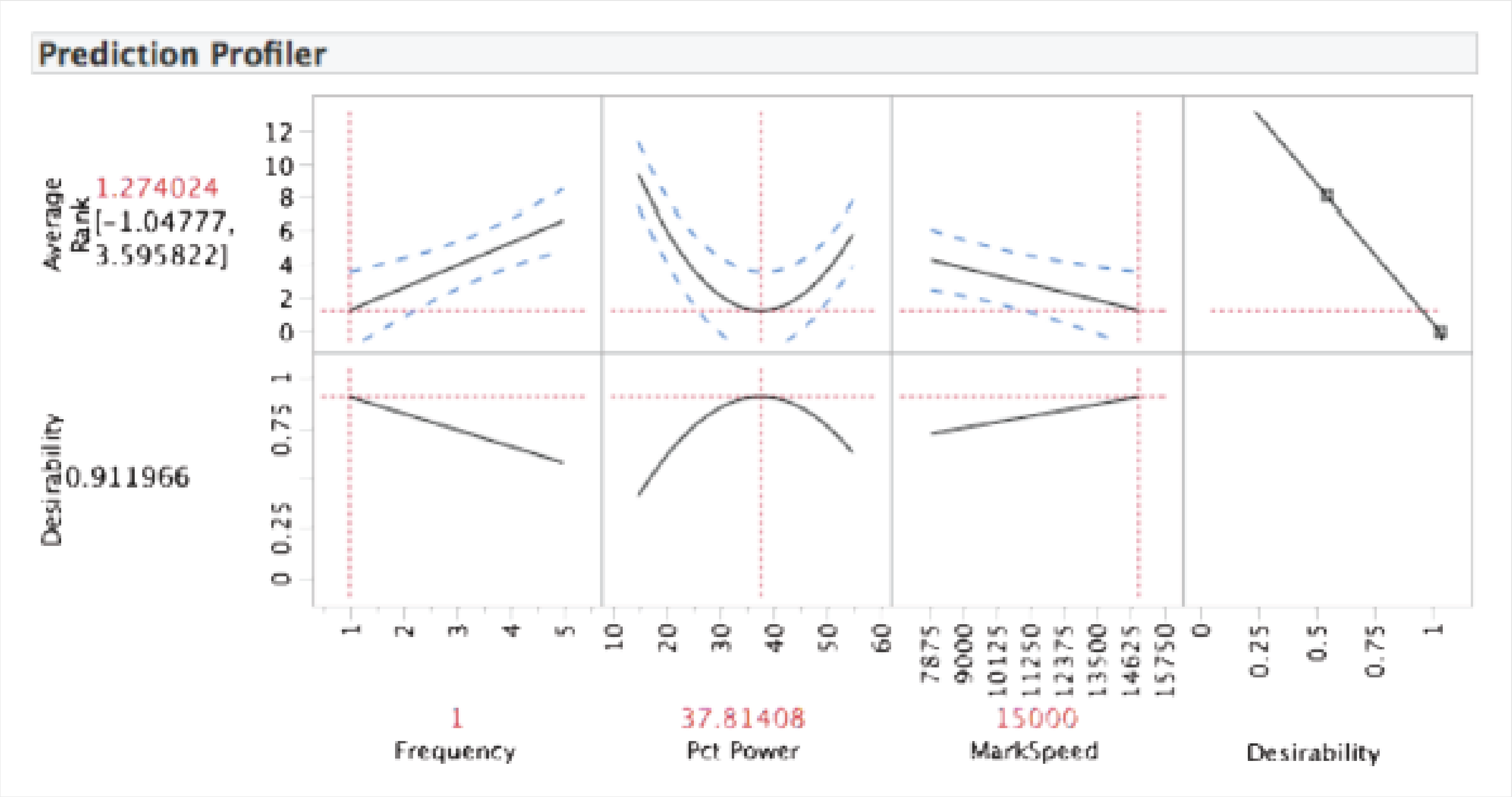
Laser Etch Design: 15-run DSD in 3 blocks

Run	MarkSpeed	Frequency	Pct Power	Repetitions	Humidity	Block
1	11500	1	15	1	Low	1
2	8000	3	15	5	High	1
3	11500	3	35	3	Medium	1
4	11500	5	55	5	High	1
5	15000	3	55	1	Low	1
6	15000	1	55	3	High	2
7	8000	5	15	3	Low	2
8	8000	1	35	1	High	2
9	11500	3	35	3	Medium	2
10	15000	5	35	5	Low	2
11	15000	1	15	5	Medium	3
12	8000	1	55	5	Low	3
13	11500	3	35	3	Medium	3
14	15000	5	15	1	High	3
15	8000	5	55	1	Medium	3

Recall: Response is average rank (smaller is better)

CARLSON
SCHOOL OF MANAGEMENT
UNIVERSITY OF MINNESOTA

Prediction profiler plots---JMP Demo



Verification run: Clear, easy to read, cheap

So, what are the current best practices?

Current best practices are:

- Problem-driven
- Three-level, where possible
- Multi-objective
- Supersaturated