CREATING AND ANALYZING DEFINITIVE SCREENING DESIGNS

Mastering JMP Webcast
February 11, 2021

Tom Donnelly, PhD, CAP
Principal Systems Engineer
JMP Defense & Aerospace Team

AGENDA

- Why do we use Design of Experiments (DOE)?
- Review of Classic DOE
- Custom DOE is all about

  Making Designs Fit the Problem –

  NOT Making Problems Fit the Designs!
- However, use Definitive Screening Designs (DSDs) – when possible!
- **Quick example of creating and fitting a DSD.**
- What are DSDs?
- How do we fit models for DSDs?
- When results are ambiguous, it is easy to augment DSD to RSM.
- Examples:
  - Extraction 3 Data.jmp : continuous with a blocking factor, & 4 extra runs
  - CO2_Process.jmp : all continuous factors, no extra runs
  - Peanut Data.jmp : continuous & categorical factors, & 4 extra runs

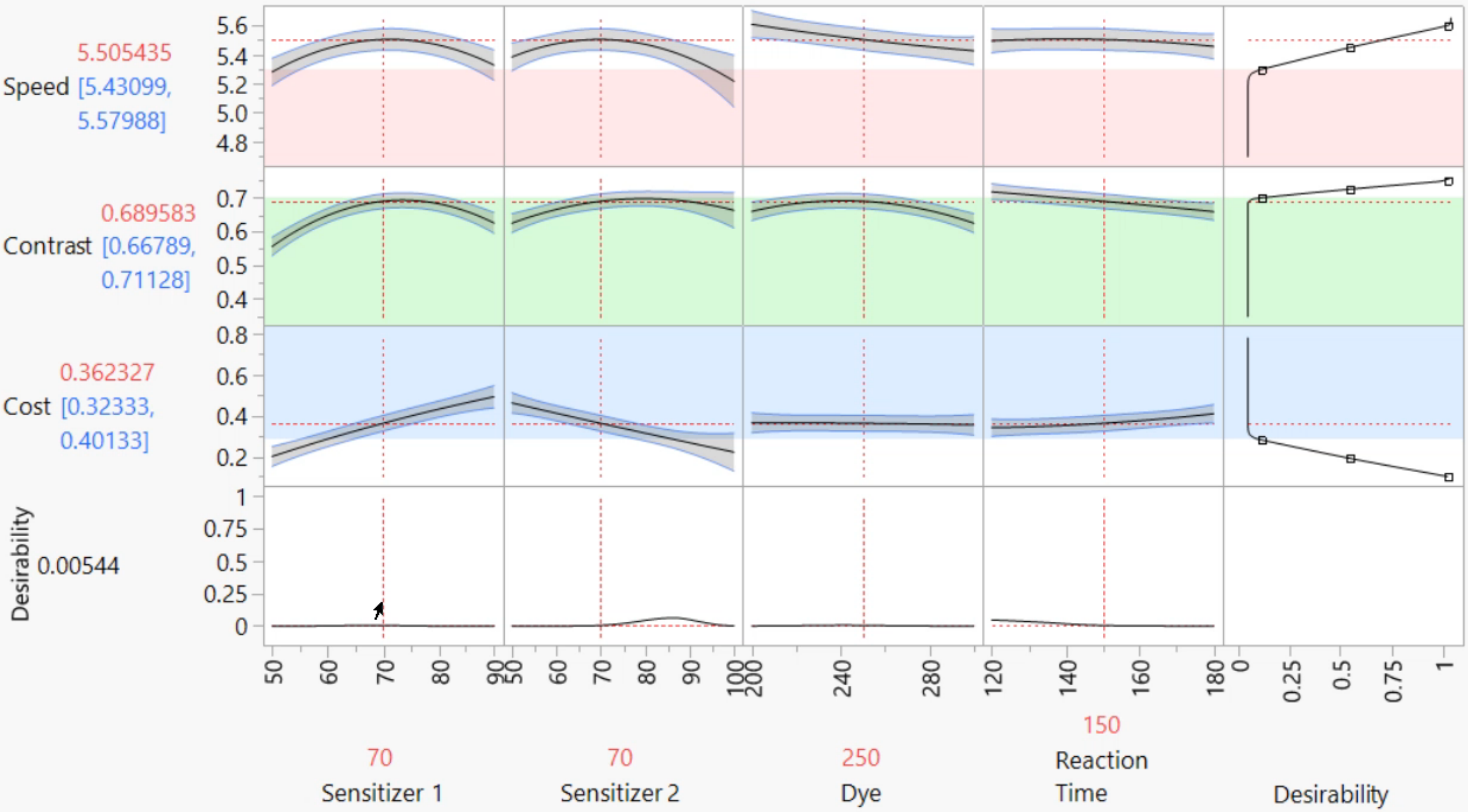
WHY USE DOE?

QUICKER ANSWERS,
LOWER COSTS,
SOLVE BIGGER PROBLEMS

- More rapidly answer "what if?" questions
- Do sensitivity and trade-space analysis
- Optimize across multiple responses
- By running efficient subsets of all possible combinations, one can – for the same resources and constraints – ***solve bigger problems***
- By running sequences of designs one can be as ***cost effective as possible*** and ***run no more trials than needed*** to get a useful answer

TRADE-OFF AND OPTIMIZATION

Prediction Profiler

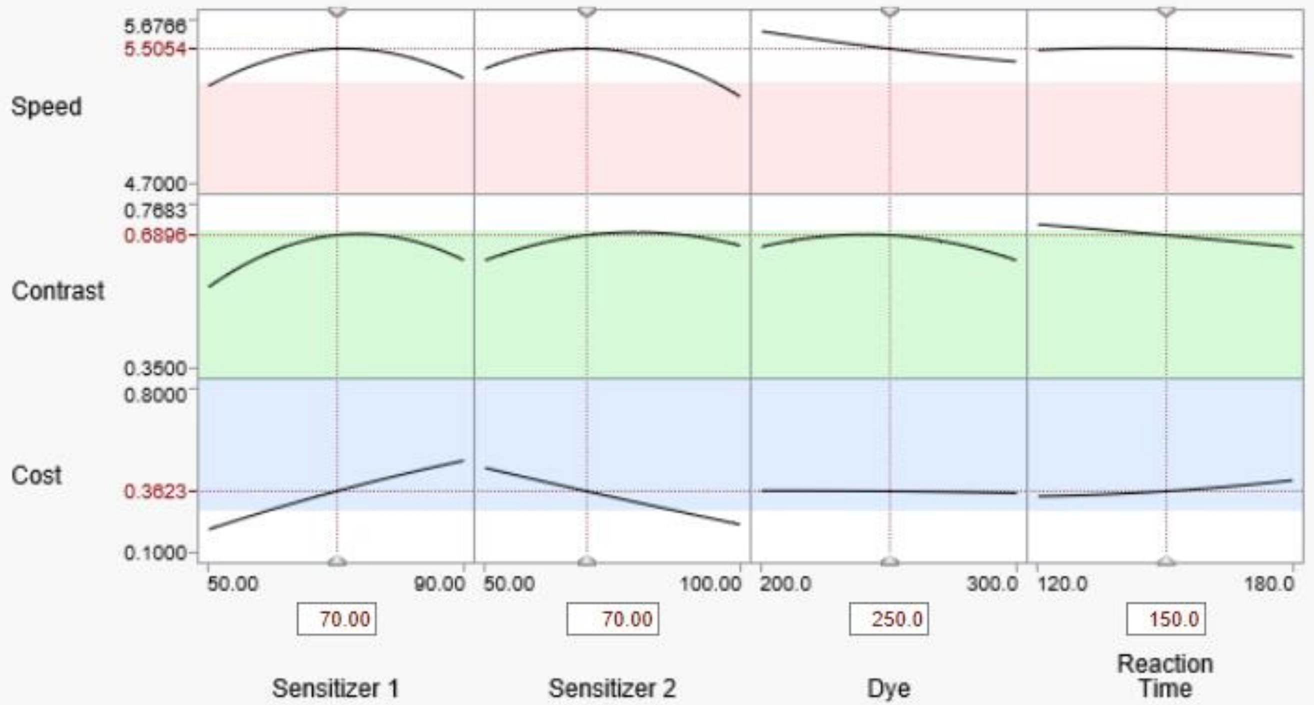


Remembered Settings

Setting	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost	Desirability
Equal Importance Opt	80.753574	91.269729	250.57625	120	5.3542877	0.7466933	0.2504014	0.347702
Mid Point Settings	70	70	250	150	5.5054353	0.6895831	0.3623274	0.004875
Cost 6X Speed & Contrast	84.016038	93.725925	283.02514	120	5.2902084	0.72549	0.1991539	0.214425
Opt Spd3X-Cntr1X-Cost6X	81.958309	90.706277	286.82246	120	5.3269582	0.7177857	0.2211116	0.264298

SHARE RESULTS ON JMP PUBLIC OR JMP LIVE

Prediction Profiler

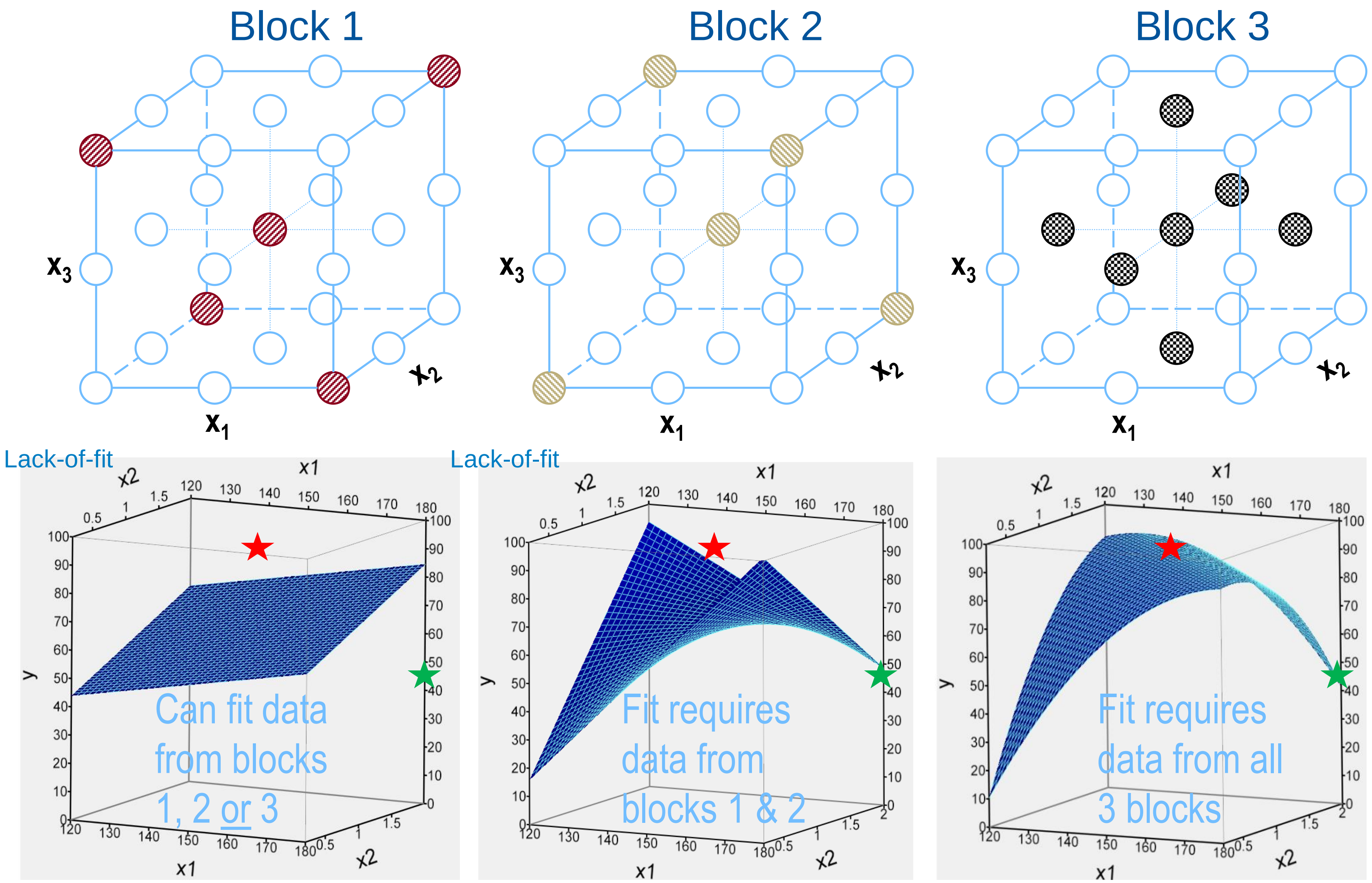


View optimizations on your phone. Scan the QR code to launch browser, then use finger to interact with the Prediction Profiler and to "Apply" saved settings.

Remembered Settings

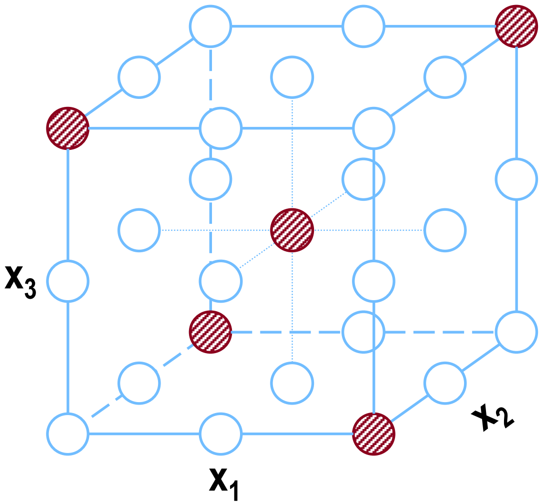
	Setting	Sensitizer 1	Sensitizer 2	Dye	Reaction Time	Speed	Contrast	Cost
Apply	Equal Importance Opt	80.753574	91.269729	250.57625	120	5.3542877	0.7466933	0.2504014
Apply	Mid Point Settings	70	70	250	150	5.5054353	0.6895831	0.3623274
Apply	Cost 6X Speed & Contrast	84.016038	93.725925	283.02514	120	5.2902084	0.72549	0.1991539
Apply	Opt Spd3X-Cntr1X-Cost6X	81.958309	90.706277	286.82246	120	5.3269582	0.7177857	0.2211116

CLASSIC RESPONSE-SURFACE DOE IN A NUTSHELL



POLYNOMIAL MODELS USED TO CALCULATE SURFACES

Block 1

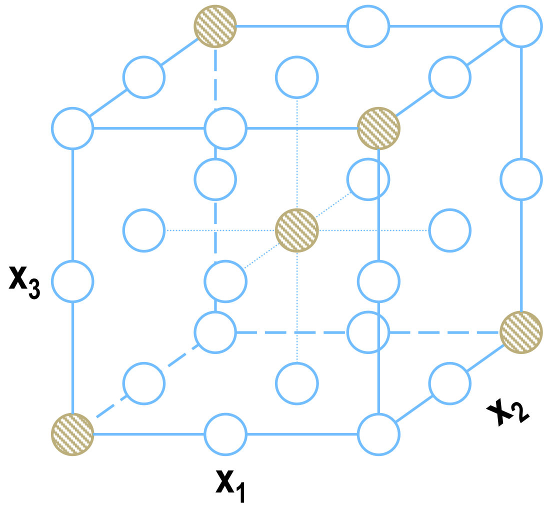


$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$

Run this block 1st to:

(i) estimate the main effects*
(ii) use center point to check for curvature.

Block 2

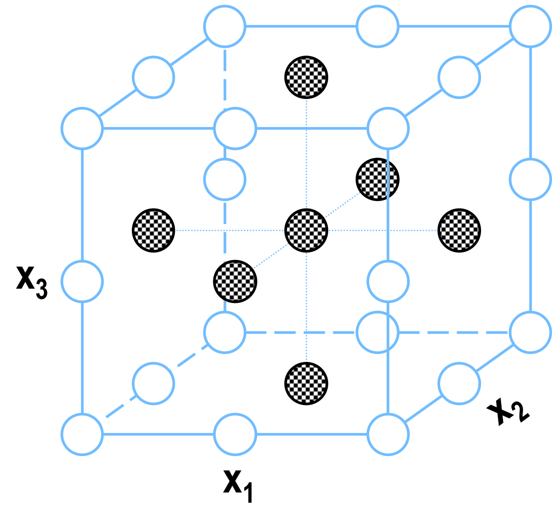


$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$
$$+ a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$

Run this block 2nd to:

(i) repeat main effects estimate,
(ii) check if process has shifted
(iii) add interaction effects to model <u>if needed.</u>

Block 3

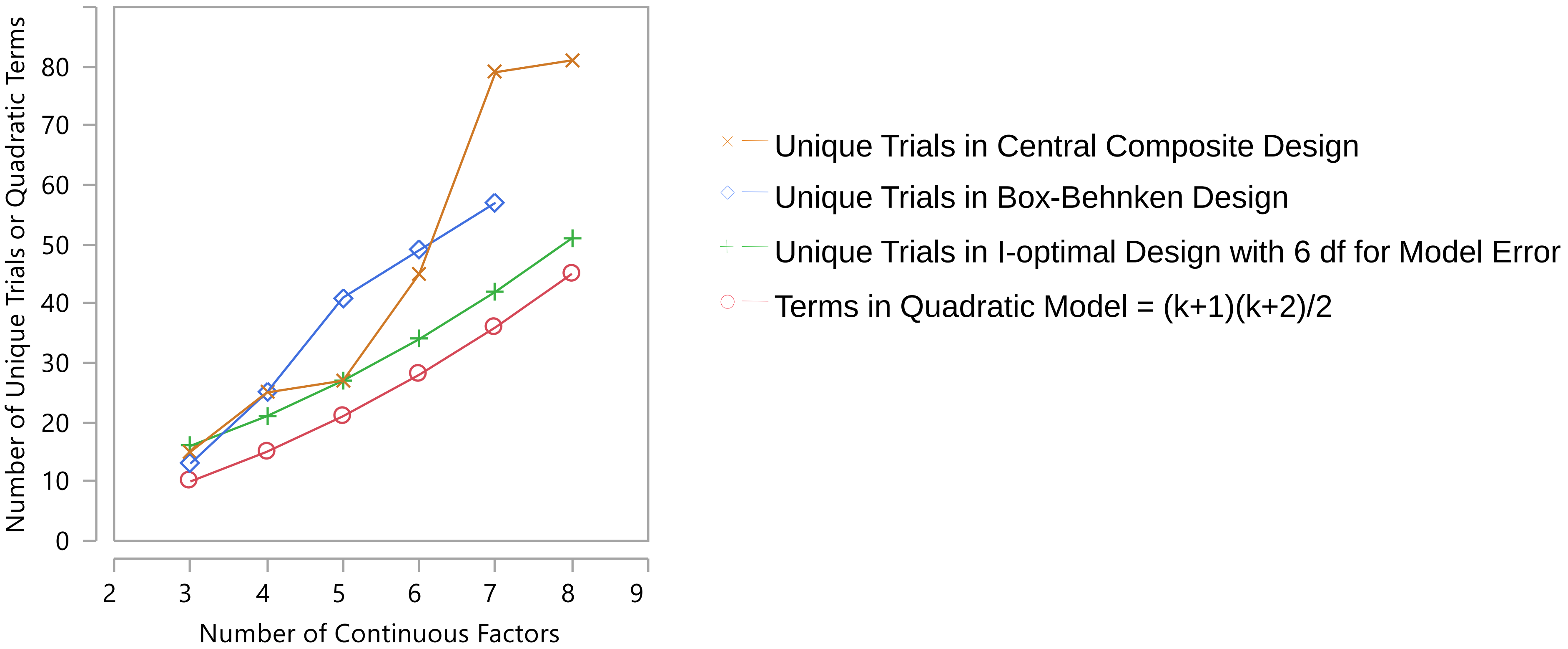


$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$
$$+ a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$
$$+ a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

Run this block 3rd to:

(i) repeat main effects estimate,
(ii) check if process has shifted
(iii) add curvature effects to model <u>if needed.</u>

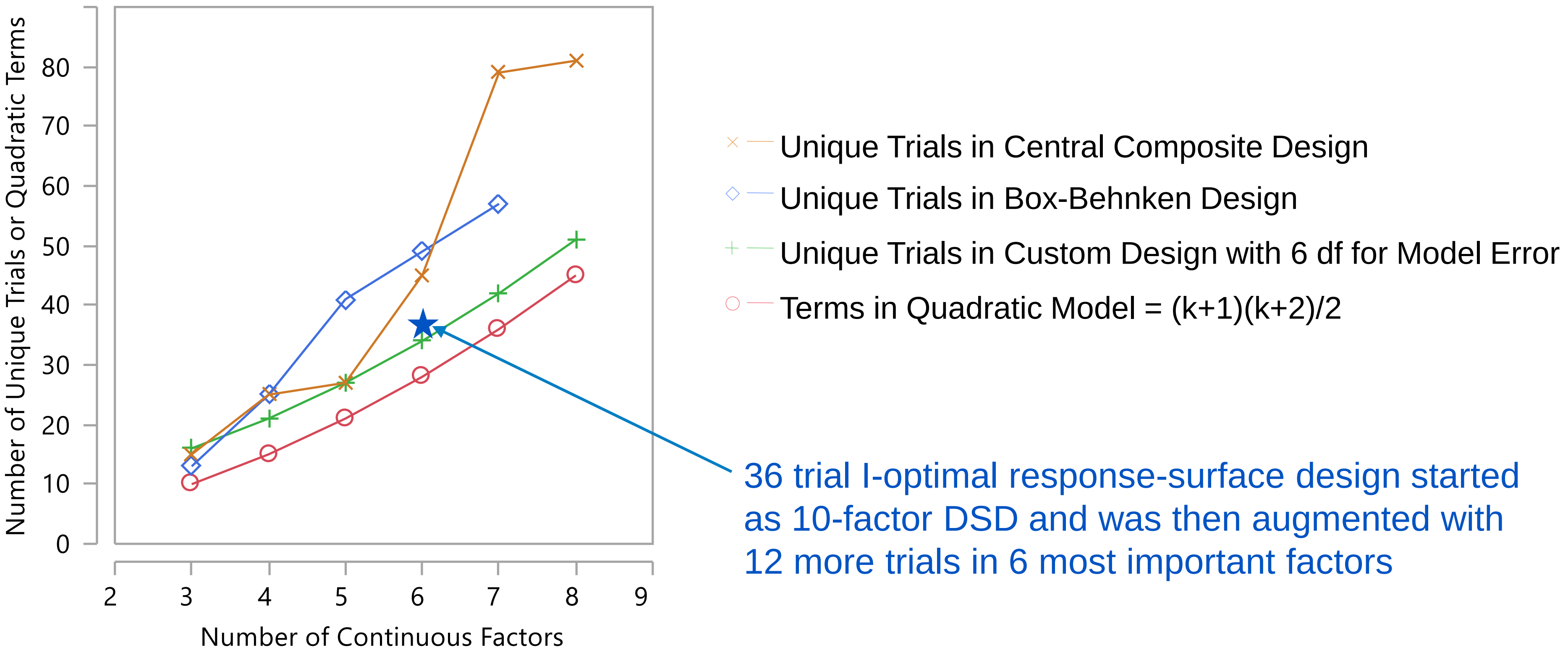
NUMBER OF UNIQUE TRIALS FOR 3 RESPONSE-SURFACE DESIGNS AND NUMBER OF QUADRATIC MODEL TERMS VS. NUMBER OF CONTINUOUS FACTORS



If generally running 3, 4 or 5-factor fractional-factorial designs…

1. How many interactions are you not investigating?
2. How many more trials needed to fit curvature?

NUMBER OF UNIQUE TRIALS FOR 3 RESPONSE-SURFACE DESIGNS AND NUMBER OF QUADRATIC MODEL TERMS VS. NUMBER OF CONTINUOUS FACTORS

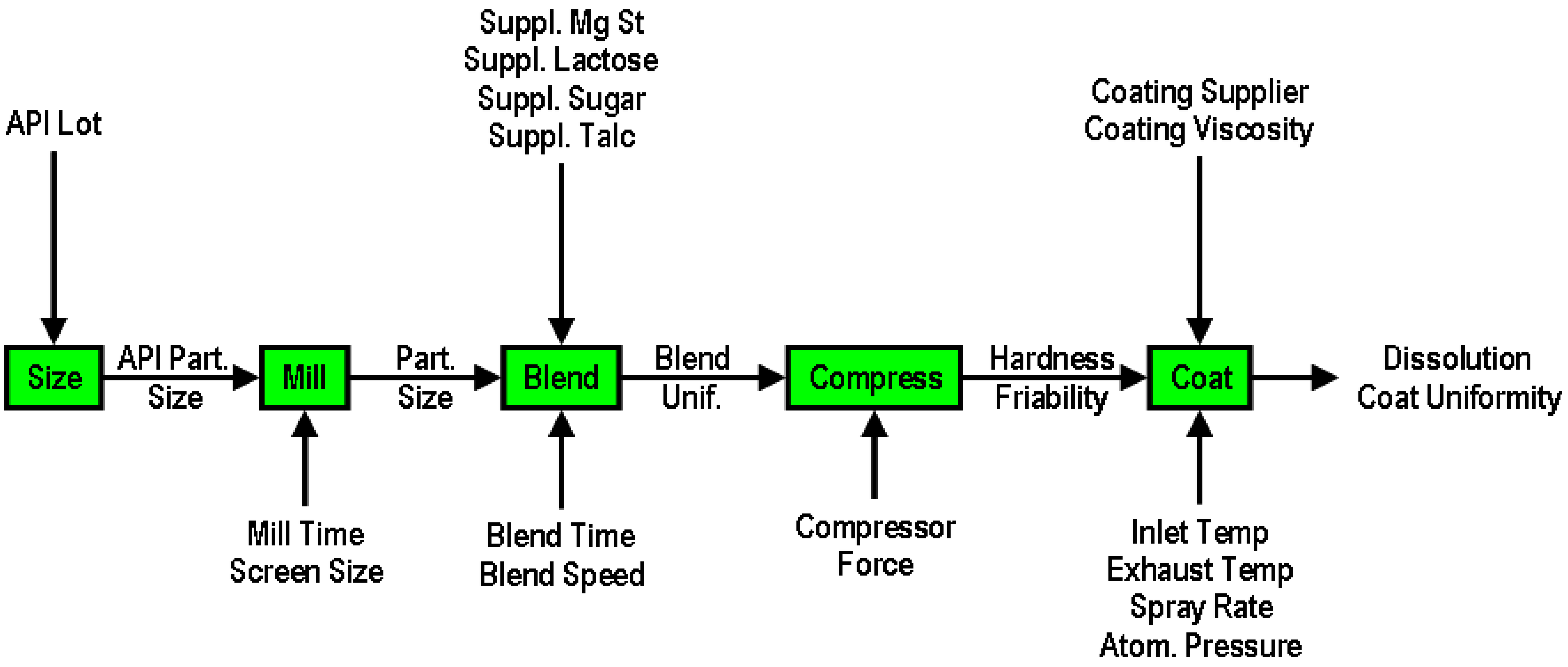


If generally running 3, 4 or 5-factor fractional-factorial designs…

1. How many interactions are you not investigating?
2. How many more trials needed to fit curvature?
3. Consider two stages: Definitive Screening + Augmentation

CLASSIC DEFINITION OF DOE

Purposeful control of the inputs (factors) in such a way as to deduce their relationships (if any) with the output (responses).



ALTERNATIVE DEFINITION OF DOE

A DOE is the specific collection of trials run to support a proposed model.

- If proposed model is ***simple***, e.g. just main effects or ***1st order*** effects (x_1, x_2, x_3, etc.), the design is called a ***screening*** DOE
  - Goals include **rank factor importance** or **find a "winner" quickly**
  - Used with many (> 6?) factors at start of process characterization
- If the proposed model is ***more complex***, e.g. the model is ***2nd order*** so that it includes two-way interaction terms (x_1x_2, x_1x_3, x_2x_3, etc.) and in the case of continuous factors, squared terms (x_1^2, x_2^2, x_3^2, etc.), the design is called a ***response-surface*** DOE
  - Goal is generally to develop a **predictive model** of the process
  - Used with a few (< 6?) factors after a screening DOE

ALTERNATIVE DEFINITION OF DOE

A DOE is the specific collection of trials run to support a proposed model.

- If proposed model is ***simple***, e.g. just main effects or ***1st order*** effects (x_1, x_2, x_3, etc.), the design is called a ***screening*** DOE
  - » Goals include **rank factor importance** or **find a "winner" quickly**
  - » Used with many (> 6?) factors at start of process characterization
- If the proposed model is ***more complex***, e.g. the model is ***2nd order*** so that it includes two-way interaction terms (x_1x_2, x_1x_3, x_2x_3, etc.) and in the case of continuous factors, squared terms (x_1^2, x_2^2, x_3^2, etc.), the design is called a ***response-surface*** DOE
  - » Goal is generally to develop a **predictive model** of the process
  - » Used with a few (< 6?) factors after a screening DOE

Definitive Screening Designs allow the fitting of second order terms – ALL squared and potentially SOME interaction terms – for no more work than classic screening designs.

REAL-WORLD DESIGN ISSUES

How many experimenters have any of these issues?
Most of these are NOT well treated by classic DOE

- Work with these different kinds of control variables/factors:
  - **Continuous/quantitative?** (Finely adjustable like *temperature*, *speed*, *force*)
  - **Categorical/qualitative?** (Comes in types, like material = *rubber*, *polycarbonate*, *steel* with mixed # of levels; 3 chemical agents, 4 decontaminants, 8 coupon materials…)
  - **Mixture/formulation?** (Blend different amounts of *ingredients* and the process performance is dependent on the *proportions* more than on the amounts)
  - **Blocking?** (e.g. "lots" of the same raw materials, multiple "same" machines, samples get processed in "groups" – like "eight in a tray," run tests over multiple days – i.e. variables for which there *shouldn't* be a causal effect
- Work with **combinations of these four kinds** of variables?
- Certain **combinations cannot be run**? (too costly, unsafe, breaks the process)
- Certain factors are **hard-to-change** (temperature takes a day to stabilize)
- Would like to **add onto existing trials**? (really expensive/time consuming to run)

REAL-WORLD DESIGN ISSUES

How many experimenters have any of these issues?
Most of these are NOT well treated by classic DOE

- Work with these different kinds of control variables/factors:
  - **Continuous/quantitative?** (Finely adjustable like *temperature*, *speed*, *force*)
  - **Categorical/qualitative?** (Comes in types, like material = *rubber*, *polycarbonate*, *steel* with mixed # of levels; 3 chemical agents, 4 decontaminants, 8 coupon materials…)
  - **Mixture/formulation?** (Blend different amounts of *ingredients* and the process performance is dependent on the *proportions* more than on the amounts)
  - **Blocking?** (e.g. "lots" of the same raw materials, multiple "same" machines, samples get processed in "groups" – like "eight in a tray," run tests over multiple days – i.e. variables for which there *shouldn't* be a causal effect
- Work with **combinations of these four kinds** of variables?
- Certain **combinations cannot be run**? (too costly, unsafe, breaks the process)
- Certain factors are **hard-to-change** (temperature takes a day to stabilize)
- Would like to **add onto existing trials**? (really expensive/time consuming to run)

Many of these issues prevent the use of Definitive Screening Designs.
BUT, if your factors are **continuous**, **2-level categorical**, and/or **blocking**
then consider doing a DSD first.

QUICK EXAMPLE FROM DOE GUIDE

Extraction 3 Data.jmp

- Uses 6 continuous factors plus blocking at 2 levels
- Add 4 extra runs DSD
- Analyze with Fit Definitive Screening (p. 276 of DOE Guide)
- Factors and Ranges shown below

Methanol	Ethanol	Propanol	Butanol	pH	Time
0	0	0	0	6	1
10	10	10	10	9	2

SUMMARY OF MODERN SCREENING DOE

- ***Definitive <u>Screening</u> Designs***
  - Efficiently estimate main and quadratic effects for no more and **often fewer trials than traditional designs**
  - If only a few factors are important the design may collapse into a **“one-shot”** design that supports a response-surface model (RSM).
  - If many factors are important (so RSM can’t be fit) the design can be **augmented** to support an RSM
  - Case study for a **10-variable process** shows that it can be **optimized in just 23 unique trials**
    - » Visually “model” factors
    - » Fit Definitive Screening
    - » Fit All Possible Models
    - » Augment design with subset of original factors

WHAT IS THE MINIMUM # FACTORS "COLLAPSE" TO RSM

- For 6 through at least 30 factors, it is possible to estimate the parameters of any **full quadratic model** involving **3** or fewer factors with high precision.
- For 18 factors or more, they can fit full quadratic models in any **4** factors.
- For 24 factors or more, they can fit full quadratic models in any **5** factors.
- **Due to factor sparsity, one can often fit response-surface models with more factors than these minimums.**

REFERENCES

Original Research on Definitive Screening Designs

Jones, B., and C. J. Nachtsheim (2011). “A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects," *Journal of Quality Technology*, 43 pp. 1-15

Xiao, L, Lin, D. K.J., and B. Fengshan (2012). “Constructing Definitive Screening Designs Using Conference Matrices,” *Journal of Quality Technology*, 44, pp. 1-7.

Jones, B., and C. J. Nachtsheim (2013). “Definitive Screening Designs with Added Two-Level Categorical Factors,” *Journal of Quality Technology*, 45 pp. 121-129

Jones, B., and C. J. Nachtsheim (2016a). “Blocking Schemes for Definitive Screening Designs,“ *Technometrics*, 58, pp. 74-83

Jones, B., and C. J. Nachtsheim (2016b). “Effective Model Selection for Definitive Screening Designs,“ *Technometrics*, (online now) https://www.tandfonline.com/doi/full/10.1080/00401706.2016.1234979.

IN ORIGINAL 2011 JQT PAPER - DESIGN SIZE IS 2M + 1

	m = 9
1	0++++++++
2	0--------
3	+0+-+--+-
4	-0-+-++-+
5	-+0-+-+--
6	+-0+-+-++
7	--+0+---+
8	++-0-+++-
9	+-+-0++--
10	-+-+0--++
11	----+0+++
12	++++-0---
13	++--++0-+
14	--++--0+-
15	---+++-0-
16	+++---+0+
17	-++--+-+0
18	+--++-+-0
19	000000000

	m = 10
1	0++-++++-+
2	0--+----+-
3	+0-++-++--
4	-0+--+--++
5	-+0---+---
6	+-0+++-+++
7	-++0+--++-
8	+--0-++--+
9	----0++++-
10	++++0----+
11	-+-++0+-++
12	+-+--0-+--
13	++----0+++
14	--++++0---
15	++++-++0+-
16	----+--0-+
17	++--++--0-
18	--++--++0+
19	+-+-+-+-+0
20	-+-+-+-+-0
21	0000000000

	m = 11
1	0-+-----+-+
2	0+-+++++-+-
3	-0--+----++
4	+0++-++++--
5	--0++++---+
6	++0----+++-
7	---0-++-++-
8	+++0+--+--+
9	+--+0+-++++
10	-++-0-+----
11	--++-0-+-+-
12	++--+0+-+-+
13	---++-0++--
14	+++--+0--++
15	-+++--+0+++
16	+---++-0---
17	-+---+-+0-+
18	+-+++-+-0+-
19	+-----++-0+
20	-+++++--+0-
21	++-+------0
22	--+-++++++0
23	00000000000

	m = 12
1	0--+-+-+++-+
2	0++-+-+---+-
3	-0+++++++---
4	+0-------+++
5	++0-++-++-++
6	--0+--+--+--
7	+--0+-+-+--+
8	-++0-+-+-++-
9	++++0-++++++
10	----0+------
11	+-+-+0++-+-+
12	-+-+-0--+-+-
13	++++-+0----+
14	----+-0++++-
15	--+++--0--++
16	++---++0++--
17	+-++++--0++-
18	-+----++0--+
19	++-++--+-0--
20	--+--++-+0++
21	-+-++++--+0+
22	+-+----++-0-
23	+--+-+++--+0
24	-++-+---++-0
25	000000000000

DEFINITIVE SCREENING DESIGNS FROM CONFERENCE MATRICES
XIAO, BAI AND LIN (JQT, 2012)

The D-efficiency is 92.3%, higher than 89.8% for the design given in Jones and Nachtsheim (2011).

$$
\mathrm{D} = \begin{pmatrix} \mathrm{C} \\ -\mathrm{C} \\ 0 \end{pmatrix} =
\begin{array}{c}
\begin{array}{cccccccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & x_{11} & x_{12} \end{array} \\
\left(\begin{array}{cccccccccccc}
0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 0 & -1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\
1 & 1 & -1 & 0 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & -1 & 0 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\
1 & 1 & 1 & -1 & -1 & 0 & 1 & 1 & -1 & -1 & 1 & -1 \\
1 & -1 & -1 & 1 & 1 & -1 & 0 & 1 & 1 & -1 & 1 & -1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 & 0 & 1 & -1 & -1 & 1 \\
1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 0 & 1 & 1 & -1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & 0 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 & 0 & 1 \\
1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 0 \\
\hdashline
0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
-1 & 0 & 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & -1 \\
-1 & -1 & 0 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\
-1 & -1 & 1 & 0 & -1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\
-1 & -1 & 1 & 1 & 0 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\
-1 & -1 & -1 & 1 & 1 & 0 & -1 & -1 & 1 & 1 & -1 & 1 \\
-1 & 1 & 1 & -1 & -1 & 1 & 0 & -1 & -1 & 1 & -1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & 1 & 0 & -1 & 1 & 1 & -1 \\
-1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 0 & -1 & -1 & 1 \\
-1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & 0 & 1 & -1 \\
-1 & 1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & 0 & -1 \\
-1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 & 1 & 0 \\
\hdashline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right)
\end{array}
$$

http://www.newton.ac.uk/programmes/DAE/seminars/090209001.pdf

CONFERENCE MATRIX METHOD IN 2012 JQT PAPER
DESIGN SIZE IS 2M + 3 FOR ODD M
DESIGN SIZE IS 2M + 1 FOR EVEN M

7-FACTOR – DSD17

	A	B	C	D	E	F	G
1	0	1	1	1	1	1	1
2	0	-1	-1	-1	-1	-1	-1
3	1	0	-1	-1	1	-1	1
4	-1	0	1	1	-1	1	-1
5	1	1	0	-1	-1	1	-1
6	-1	-1	0	1	1	-1	1
7	1	1	1	0	-1	-1	1
8	-1	-1	-1	0	1	1	-1
9	1	-1	1	1	0	-1	-1
10	-1	1	-1	-1	0	1	1
11	1	1	-1	1	1	0	-1
12	-1	-1	1	-1	-1	0	1
13	1	-1	1	-1	1	1	0
14	-1	1	-1	1	-1	-1	0
15	1	-1	-1	1	-1	1	1
16	-1	1	1	-1	1	-1	-1
17	0	0	0	0	0	0	0

8-FACTOR – DSD17

	A	B	C	D	E	F	G	H
1	0	1	1	1	1	1	1	1
2	0	-1	-1	-1	-1	-1	-1	-1
3	1	0	-1	-1	1	-1	1	1
4	-1	0	1	1	-1	1	-1	-1
5	1	1	0	-1	-1	1	-1	1
6	-1	-1	0	1	1	-1	1	-1
7	1	1	1	0	-1	-1	1	-1
8	-1	-1	-1	0	1	1	-1	1
9	1	-1	1	1	0	-1	-1	1
10	-1	1	-1	-1	0	1	1	-1
11	1	1	-1	1	1	0	-1	-1
12	-1	-1	1	-1	-1	0	1	1
13	1	-1	1	-1	1	1	0	-1
14	-1	1	-1	1	-1	-1	0	1
15	1	-1	-1	1	-1	1	1	0
16	-1	1	1	-1	1	-1	-1	0
17	0	0	0	0	0	0	0	0

Both designs are orthogonal in linear and squared terms
Factor H will become a hidden Fake Factor in DSD Analysis

DEFINITIVE SCREENING DESIGNS HAVE DESIRABLE PROPERTIES

- Main effects are not confounded with 2^{nd} order effects
- Number of trials for even numbers of factors is (2m + 1)

  and for odd numbers of factors it is (2m + 3)

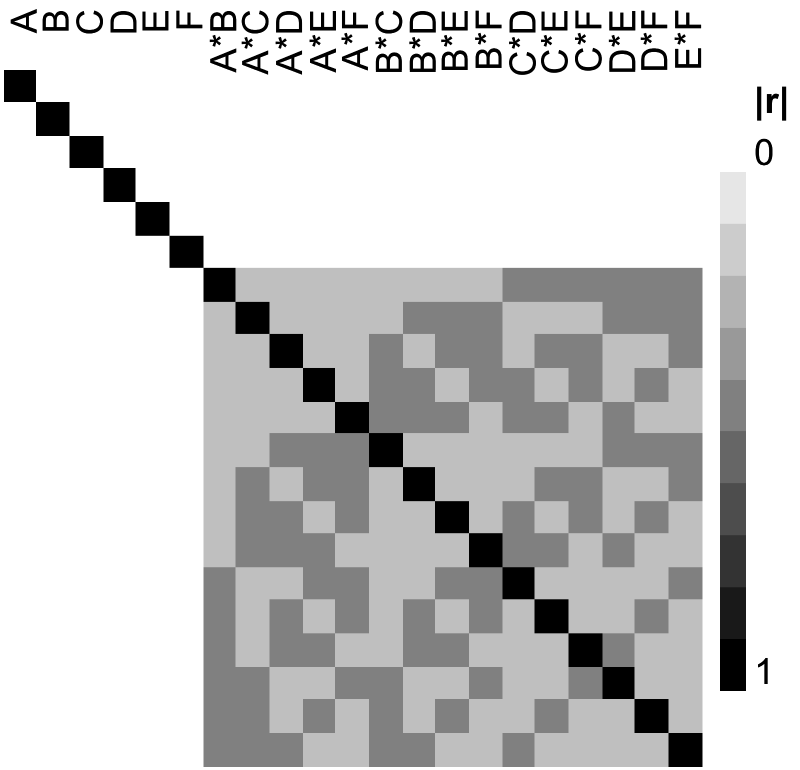
  which is **equal to or smaller** than a Plackett-Burman (Res III) or Fractional Factorial (Res IV) design plus center point
- There are mid-levels for each factor allowing estimation of **curvature individually - not just globally** as with a PB or FF designs plus center point
- If drop a factor, the design retains all its properties
- If a subset of factors are significant there is a good chance that interaction terms may also be fit

  The screening design may even collapse into a response-surface design supporting a 2^{nd} order model in a subset of factors with which one can optimize the process

6-FACTOR, 13-TRIAL, DEFINITIVE SCREENING DESIGN

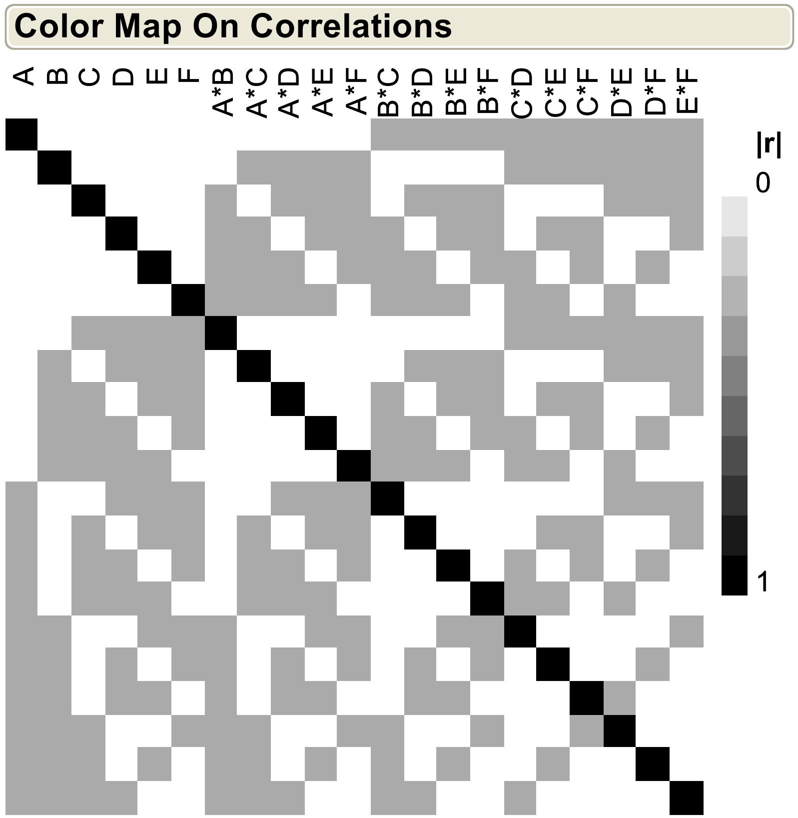
	A	B	C	D	E	F
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0
13	0	0	0	0	0	0

Color Map On Correlations

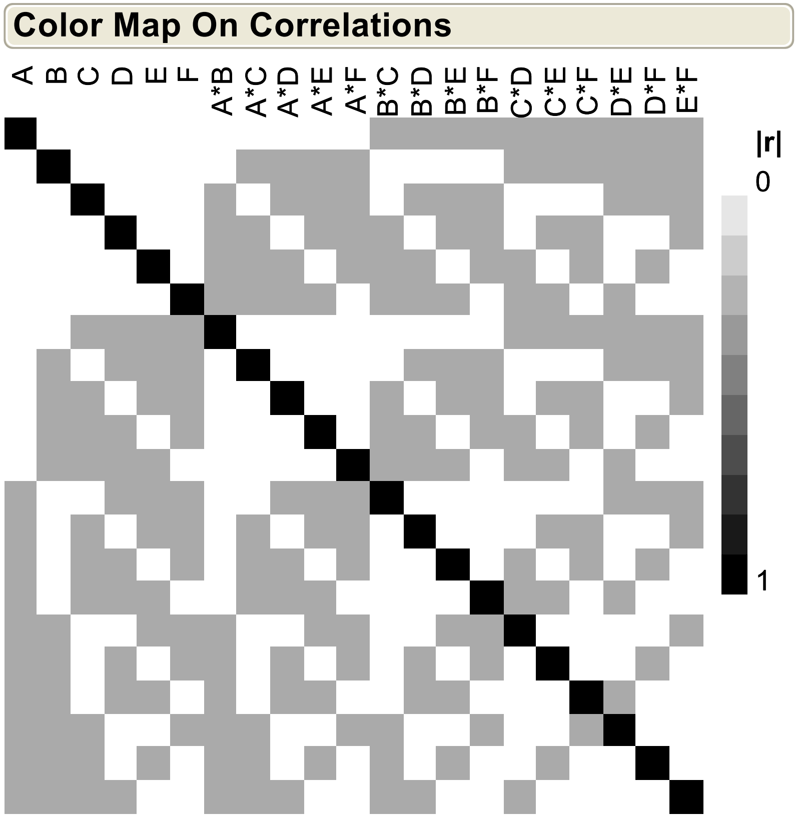


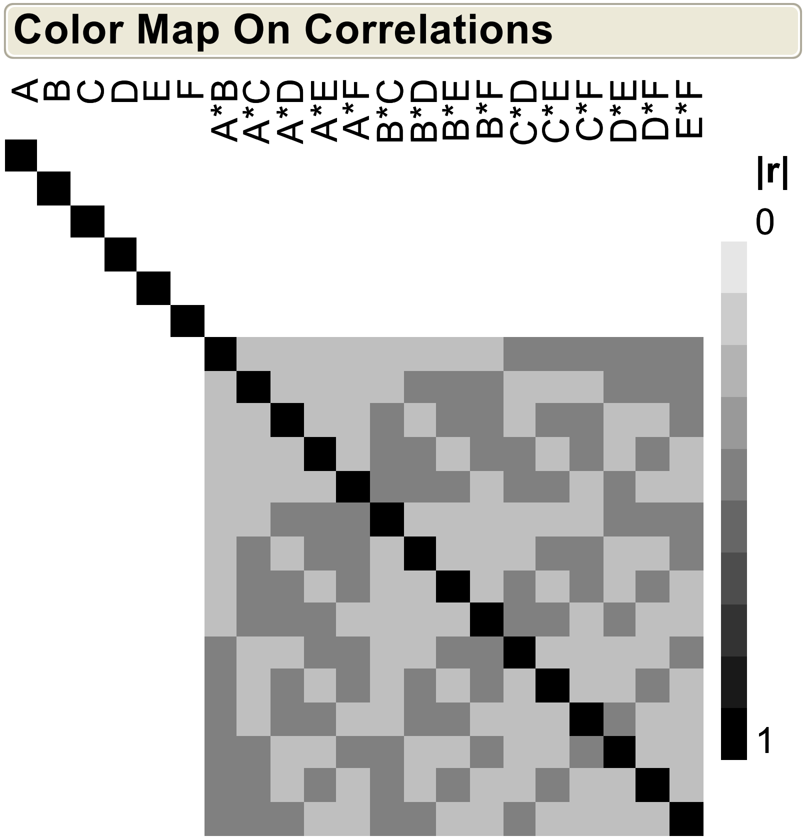
6-FACTOR, 12-TRIAL, PLACKETT-BURMAN DESIGN

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



COLOR MAPS FOR 6-FACTOR, PLACKETT-BURMAN (LEFT) AND DEFINITIVE SCREENING DESIGN (RIGHT)



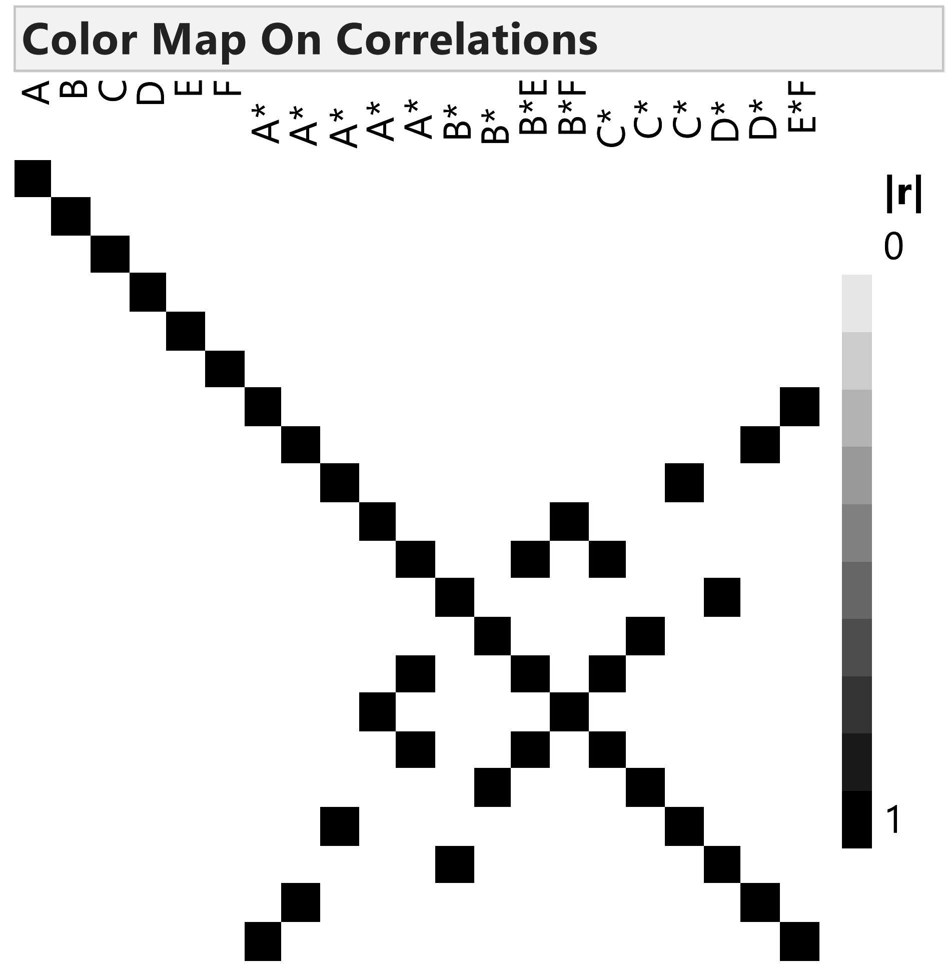


Including center point with Plackett-Burman, these two designs are both 13 trials

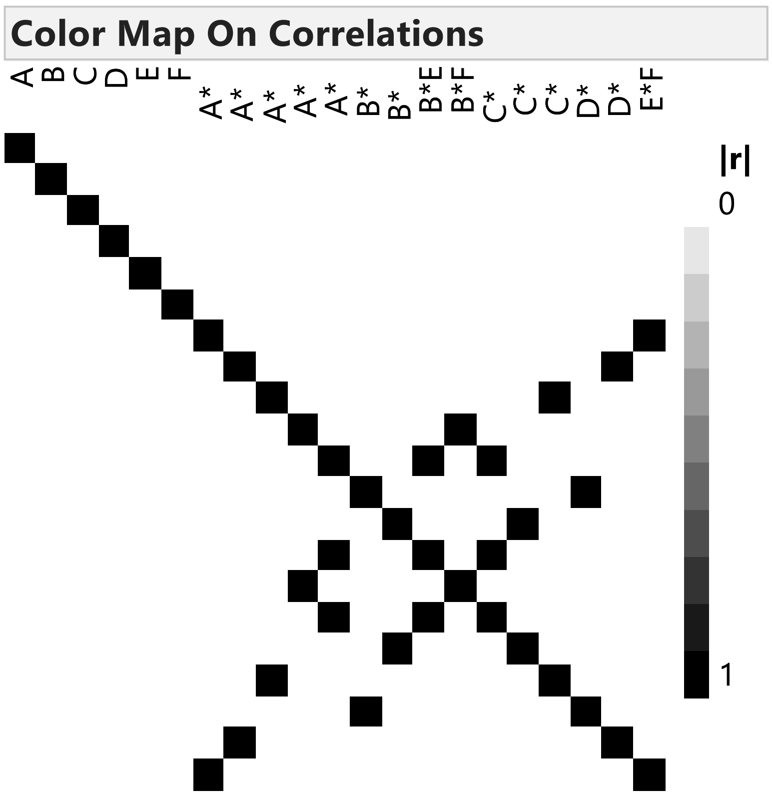
Same size BUT Definitive Screening can test for curvature in each factor

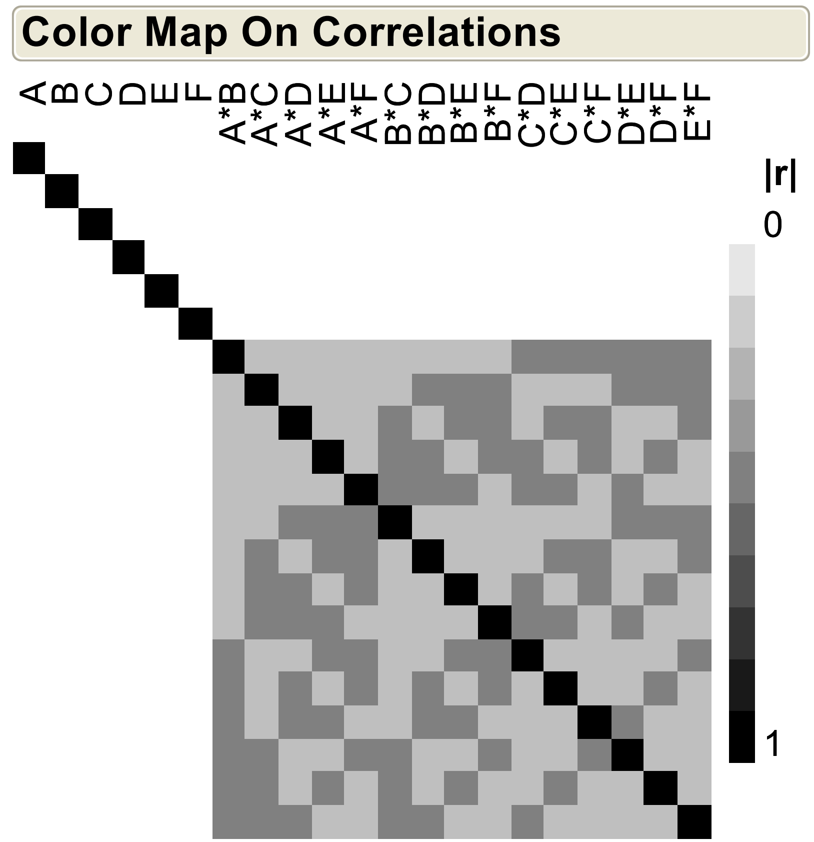
6-FACTOR, 16-TRIAL, REGULAR FRACTIONAL FACTORIAL

	Pattern	A	B	C	D	E	F
1	−−−−−−	-1	-1	-1	-1	-1	-1
2	−−−+++	-1	-1	-1	1	1	1
3	−−+−++	-1	-1	1	-1	1	1
4	−−++−−	-1	-1	1	1	-1	-1
5	−+−−+−	-1	1	-1	-1	1	-1
6	−+−+−+	-1	1	-1	1	-1	1
7	−++−−+	-1	1	1	-1	-1	1
8	−++++−	-1	1	1	1	1	-1
9	+−−−−+	1	-1	-1	-1	-1	1
10	+−−++−	1	-1	-1	1	1	-1
11	+−+−+−	1	-1	1	-1	1	-1
12	+−++−+	1	-1	1	1	-1	1
13	++−−++	1	1	-1	-1	1	1
14	++−+−−	1	1	-1	1	-1	-1
15	+++−−−	1	1	1	-1	-1	-1
16	++++++	1	1	1	1	1	1



COLOR MAPS FOR 6-FACTOR, FRACTIONAL FACTORIAL (LEFT) AND DEFINITIVE SCREENING DESIGN (RIGHT)





Including center point with FF increases size to 17 trials - 13-trial Definitive Screening Design is **4 fewer tests AND can test for curvature in each factor**
Or, add 4 extra rows to DSD to improve robustness of Fitting Models

DO WE GIVE UP NOTHING?

- Relative to same size classic 2-level screening designs
  - Confidence intervals increase – typically ≤10%
  - Standard error increases – typically ≤ 10%
  - Power is reduced for main effects – typically ≤ 10% (comparing just ME)
  - Power for squared terms is “low”
    - Still better than power for single center point test for curvature
    - Power is same as larger Central Composite Design supporting full quadratic model
    - Power increases as fewer curvature terms are evaluated – drop least important terms (Factor Sparsity is our friend!)

ANY OTHER WEAKNESSES?

- Factor range for screening may not include optimum
  - So, follow on design will be over different ranges – really can’t augment
  - This is more likely with early product development than with designs testing mature systems

CONFIDENCE INTERVAL, STANDARD ERROR & MAIN EFFECTS POWER FOR 6-FACTOR DESIGNS:

PLACKETT-BURMAN 12 + CP
DEFINITIVE SCREENING DESIGN 13
FRACTIONAL-FACTORIAL 16 + CP
DEFINITIVE SCREENING DESIGN 17

PB12+CP

Estimation Efficiency

Parameter	Fractional Increase in CI Length	Relative Std Error of Parameters
Intercept	0	0.277
X1	0.041	0.289
X2	0.041	0.289
X3	0.041	0.289
X4	0.041	0.289
X5	0.041	0.289
X6	0.041	0.289

Power Analysis

Significance Level 0.05
Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.85
X1	1	0.821
X2	1	0.821
X3	1	0.821
X4	1	0.821
X5	1	0.821
X6	1	0.821

FF16+CP

Estimation Efficiency

Parameter	Fractional Increase in CI Length	Relative Std Error of Parameters
Intercept	0	0.243
X1	0.031	0.25
X2	0.031	0.25
X3	0.031	0.25
X4	0.031	0.25
X5	0.031	0.25
X6	0.031	0.25

Power Analysis

Significance Level 0.05
Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.959
X1	1	0.949
X2	1	0.949
X3	1	0.949
X4	1	0.949
X5	1	0.949
X6	1	0.949

DSD13

Estimation Efficiency

Parameter	Fractional Increase in CI Length	Relative Std Error of Parameters
Intercept	0	0.277
X1	0.14	0.316
X2	0.14	0.316
X3	0.14	0.316
X4	0.14	0.316
X5	0.14	0.316
X6	0.14	0.316

+ 10% + 9%

Power Analysis

Significance Level 0.05
Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.85
X1	1	0.75
X2	1	0.75
X3	1	0.75
X4	1	0.75
X5	1	0.75
X6	1	0.75

- 9%

DSD17

Estimation Efficiency

Parameter	Fractional Increase in CI Length	Relative Std Error of Parameters
Intercept	0	0.243
X1	0.102	0.267
X2	0.102	0.267
X3	0.102	0.267
X4	0.102	0.267
X5	0.102	0.267
X6	0.102	0.267

+ 7% + 7%

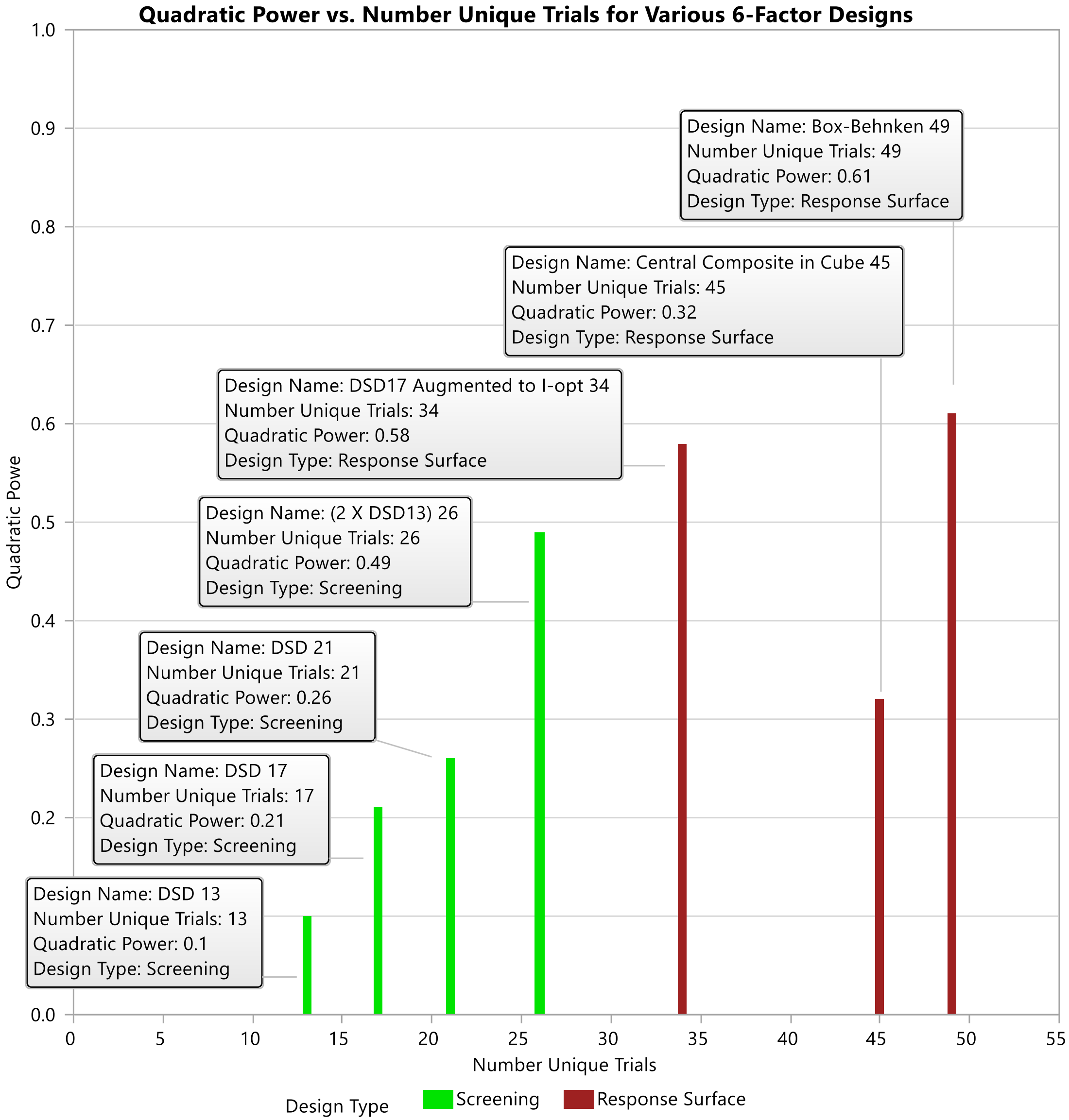
Power Analysis

Significance Level 0.05
Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.959
X1	1	0.92
X2	1	0.92
X3	1	0.92
X4	1	0.92
X5	1	0.92
X6	1	0.92

- 3%

QUADRATIC TERM POWER FOR 6-FACTOR DESIGNS – SCREENING & RSM



QUADRATIC TERM POWER FOR TEN 6-FACTOR DESIGNS – SCREENING & RSM

DSD13

0.10

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.073
X1	1	0.196
X2	1	0.196
X3	1	0.196
X4	1	0.196
X5	1	0.196
X6	1	0.196
X1*X1	1	0.096
X2*X2	-1	0.096
X3*X3	1	0.096
X4*X4	-1	0.096
X5*X5	1	0.096
X6*X6	-1	0.096

DSD17

0.21

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.13
X1	1	0.796
X2	1	0.796
X3	1	0.796
X4	1	0.796
X5	1	0.796
X6	1	0.796
X1*X1	1	0.211
X2*X2	-1	0.211
X3*X3	1	0.211
X4*X4	-1	0.211
X5*X5	1	0.211
X6*X6	-1	0.211

DSD21

0.26

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.159
X1	1	0.959
X2	1	0.959
X3	1	0.959
X4	1	0.959
X5	1	0.959
X6	1	0.959
X1*X1	1	0.261
X2*X2	-1	0.261
X3*X3	1	0.261
X4*X4	-1	0.261
X5*X5	1	0.261
X6*X6	-1	0.261

2X DSD13

0.49

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.259
X1	1	0.985
X2	1	0.985
X3	1	0.985
X4	1	0.985
X5	1	0.985
X6	1	0.985
X1*X1	1	0.488
X2*X2	-1	0.488
X3*X3	1	0.488
X4*X4	-1	0.488
X5*X5	1	0.488
X6*X6	-1	0.488

AUGMENT DSD17 TO I-OPT34

0.58

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.39
X1	1	0.994
X2	1	0.996
X3	1	0.996
X4	1	0.996
X5	1	0.993
X6	1	0.993
X1*X1	1	0.583
X2*X2	-1	0.587
X3*X3	1	0.568
X4*X4	-1	0.623
X5*X5	1	0.574
X6*X6	-1	0.559

PB12+CP

0.12

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.13
X1	1	0.789
X2	1	0.789
X3	1	0.789
X4	1	0.789
X5	1	0.789
X6	1	0.789
X1*X1	1	0.124

FF16+CP

0.14

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.146
X1	1	0.944
X2	1	0.944
X3	1	0.944
X4	1	0.944
X5	1	0.944
X6	1	0.944
X1*X1	1	0.14

CCD45

0.32

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.839
X1	1	1
X2	1	1
X3	1	1
X4	1	1
X5	1	1
X6	1	1
X1*X1	1	0.321
X2*X2	1	0.321
X3*X3	1	0.321
X4*X4	1	0.321
X5*X5	1	0.321
X6*X6	1	0.321

BB49

0.61

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.164
X1	1	0.997
X2	1	0.997
X3	1	0.997
X4	1	0.997
X5	1	0.997
X6	1	0.997
X1*X1	1	0.608
X2*X2	-1	0.608
X3*X3	1	0.608
X4*X4	-1	0.608
X5*X5	1	0.608
X6*X6	-1	0.608

I-OPT34

0.63

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.466
X1	1	0.995
X2	1	0.991
X3	1	0.992
X4	1	0.995
X5	1	0.989
X6	1	0.991
X1*X1	1	0.597
X2*X2	-1	0.659
X3*X3	1	0.693
X4*X4	-1	0.631
X5*X5	1	0.594
X6*X6	-1	0.621

POWER FOR 6 MAIN EFFECTS & 6 QUADRATIC TERMS FOR ALL TERMS VS. ONE QUAD TERM AT A TIME

DSD13 — **0.10**

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.073
X1	1	0.196
X2	1	0.196
X3	1	0.196
X4	1	0.196
X5	1	0.196
X6	1	0.196
X1*X1	1	0.096
X2*X2	-1	0.096
X3*X3	1	0.096
X4*X4	-1	0.096
X5*X5	1	0.096
X6*X6	-1	0.096

DSD13 — **0.24**

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.291
X1	1	0.716
X2	1	0.716
X3	1	0.716
X4	1	0.716
X5	1	0.716
X6	1	0.716
X1*X1	1	0.236

PB12+CP — **0.12**

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.13
X1	1	0.789
X2	1	0.789
X3	1	0.789
X4	1	0.789
X5	1	0.789
X6	1	0.789
X1*X1	1	0.124

DSD17 — **0.21**

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.13
X1	1	0.796
X2	1	0.796
X3	1	0.796
X4	1	0.796
X5	1	0.796
X6	1	0.796
X1*X1	1	0.211
X2*X2	-1	0.211
X3*X3	1	0.211
X4*X4	-1	0.211
X5*X5	1	0.211
X6*X6	-1	0.211

DSD17 — **0.29**

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.341
X1	1	0.913
X2	1	0.913
X3	1	0.913
X4	1	0.913
X5	1	0.913
X6	1	0.913
X1*X1	1	0.29

FF16+CP — **0.14**

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.146
X1	1	0.944
X2	1	0.944
X3	1	0.944
X4	1	0.944
X5	1	0.944
X6	1	0.944
X1*X1	1	0.14

July 22, 2010

Secretary Chu Announces Six Projects to Convert Captured CO2 Emissions from Industrial Sources into Useful Products

$106 Million Recovery Act Investment will Reduce CO2 Emissions and Mitigate Climate Change

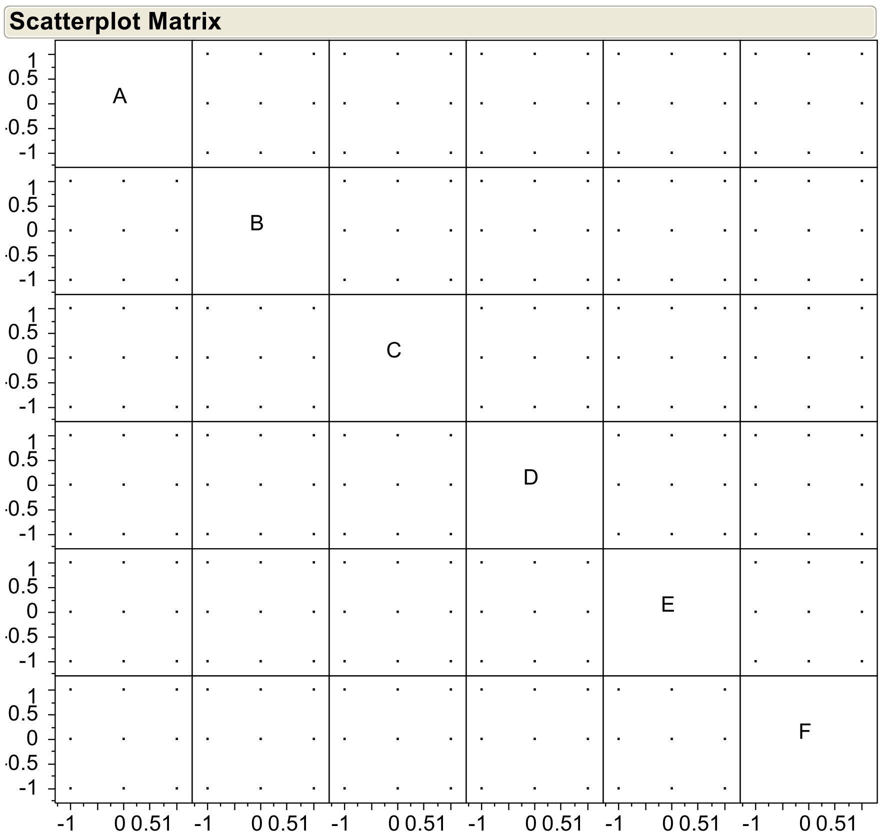
Washington, D.C. - U.S. Energy Secretary Steven Chu announced today the selections of six projects that aim to find ways of converting captured carbon dioxide (CO_2) emissions from industrial sources into useful products such as fuel, plastics, cement, and fertilizers. Funded with $106 million from the American Recovery and Reinvestment Act -matched with $156 million in private cost-share -today's selections demonstrate the potential opportunity to use CO_2 as an inexpensive raw material that can help reduce carbon dioxide emissions while producing useful by-products that Americans can use.

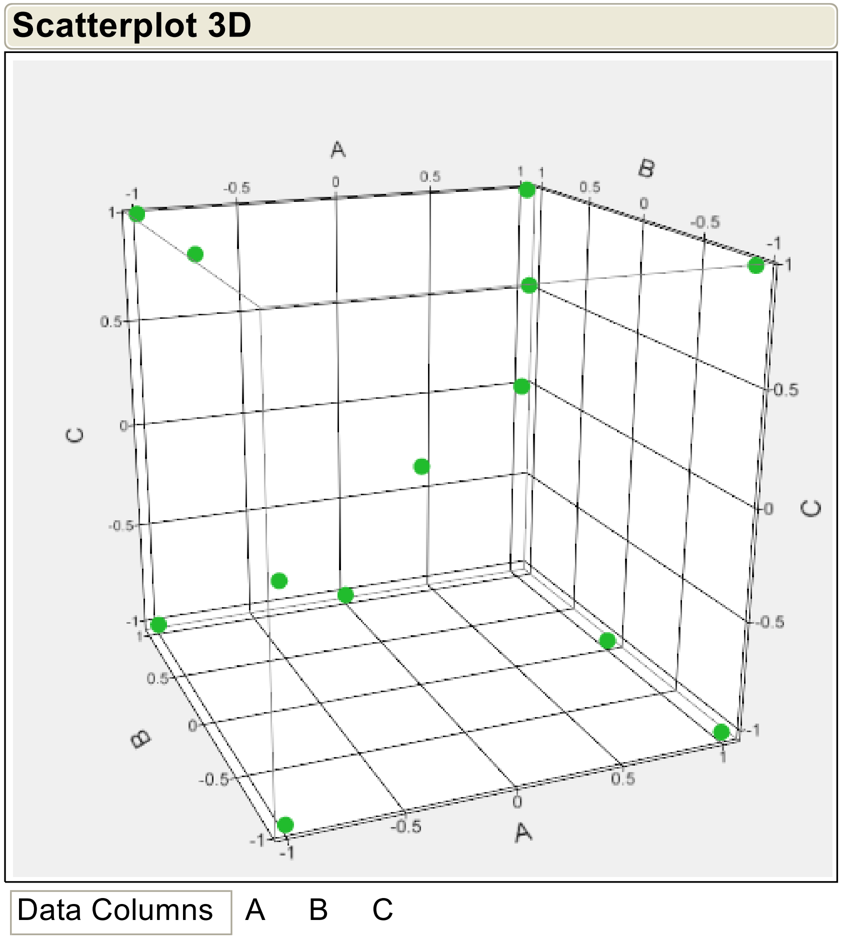
"These innovative projects convert carbon pollution from a climate threat to an economic resource," said Secretary Chu. "This is part of our broad commitment to unleash the American innovation machine and build the thriving, clean energy economy of the future."

23/1	Yield @ Time t	A	B	C	D	E	F	G	H	I	J
1	1.38	-1	1	1	0	1	-1	1	-1	1	1
2	6.44	1	-1	-1	-1	1	-1	1	1	0	1
3	5.96	-1	-1	1	-1	-1	1	-1	1	1	0
4	4.34	0	-1	1	1	1	1	1	1	-1	-1
5	10.46	-1	-1	-1	-1	-1	0	1	-1	-1	-1
6	6.95	-1	-1	1	-1	1	-1	-1	0	-1	-1
7	8.58	1	0	-1	1	1	-1	-1	-1	1	-1
8	2.69	0	1	-1	-1	-1	-1	-1	-1	1	1
9	4.3	-1	1	-1	1	0	-1	-1	1	-1	1
10	0.77	1	-1	1	-1	0	1	1	-1	1	-1
11	2.87	-1	1	1	1	-1	1	-1	-1	0	-1
12	1.01	1	1	1	1	1	0	-1	1	1	1
13	9.47	-1	-1	-1	1	1	1	0	-1	1	1
14	7.49	0	0	0	0	0	0	0	0	0	0
15	0.98	1	1	-1	1	1	-1	1	-1	-1	0
16	0.86	1	1	1	-1	-1	-1	0	1	-1	-1
17	1.25	-1	1	-1	-1	1	1	1	1	1	-1
18	1.03	1	-1	1	1	-1	-1	-1	-1	-1	1
19	1.07	1	1	0	-1	1	1	-1	-1	-1	1
20	7.33	0	0	0	0	0	0	0	0	0	0
21	2.61	1	-1	-1	0	-1	1	-1	1	-1	-1
22	11.39	-1	-1	0	1	-1	-1	1	1	1	-1
23	12.96	-1	0	1	-1	-1	1	1	1	-1	1
24	1.18	1	1	-1	1	-1	1	1	0	1	1

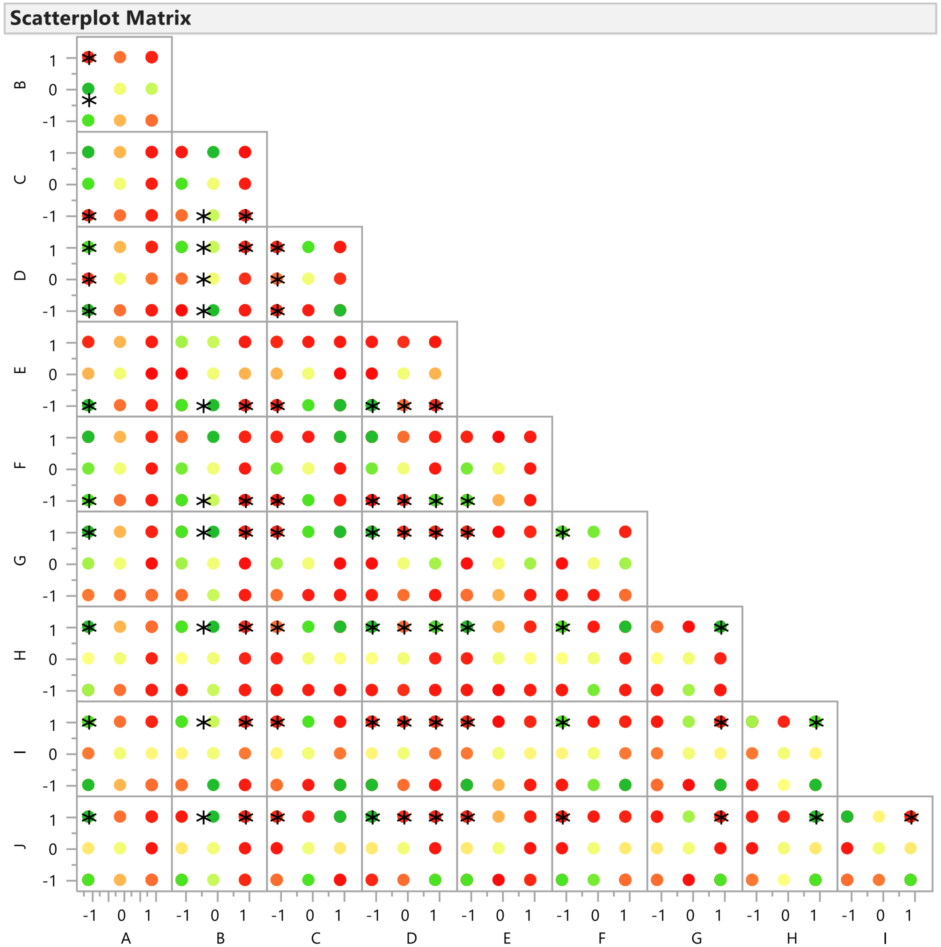
Original design was for 11 variables with 23 unique trials and the center point replicated once.

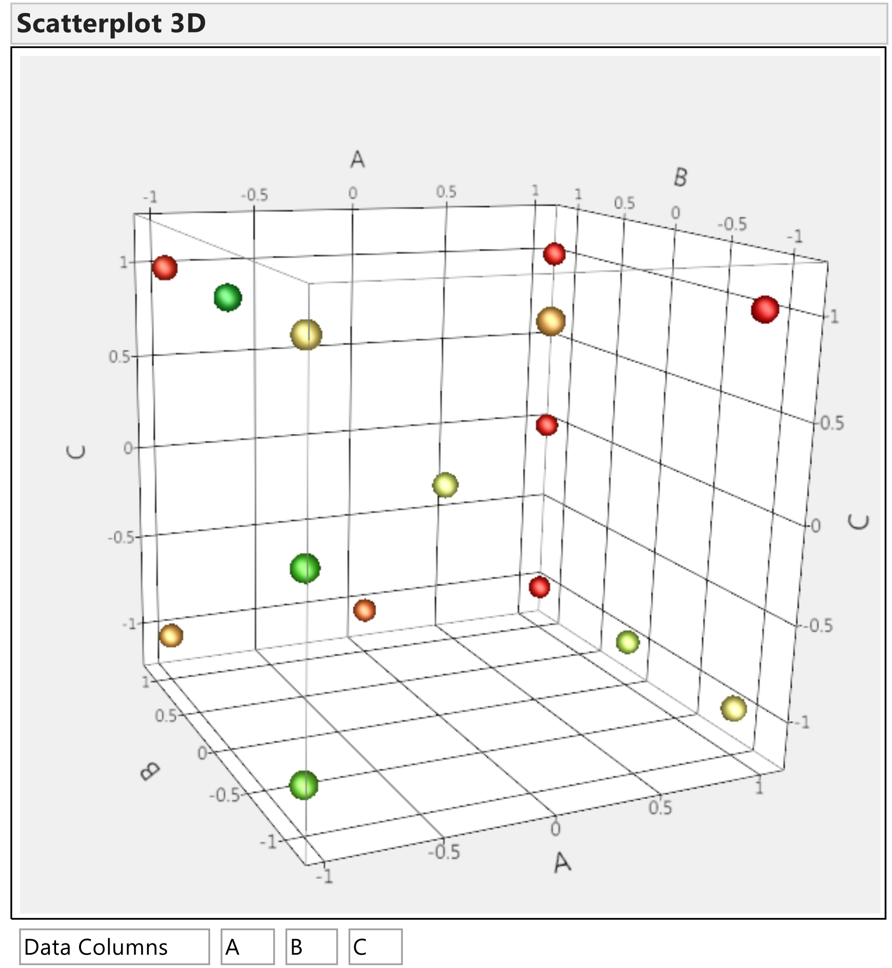
6-FACTOR DEFINITIVE SCREENING DESIGN, PROJECTION IN ALL 2-FACTOR COMBINATIONS (LEFT) AND PROJECTION IN FIRST THREE FACTORS (RIGHT)





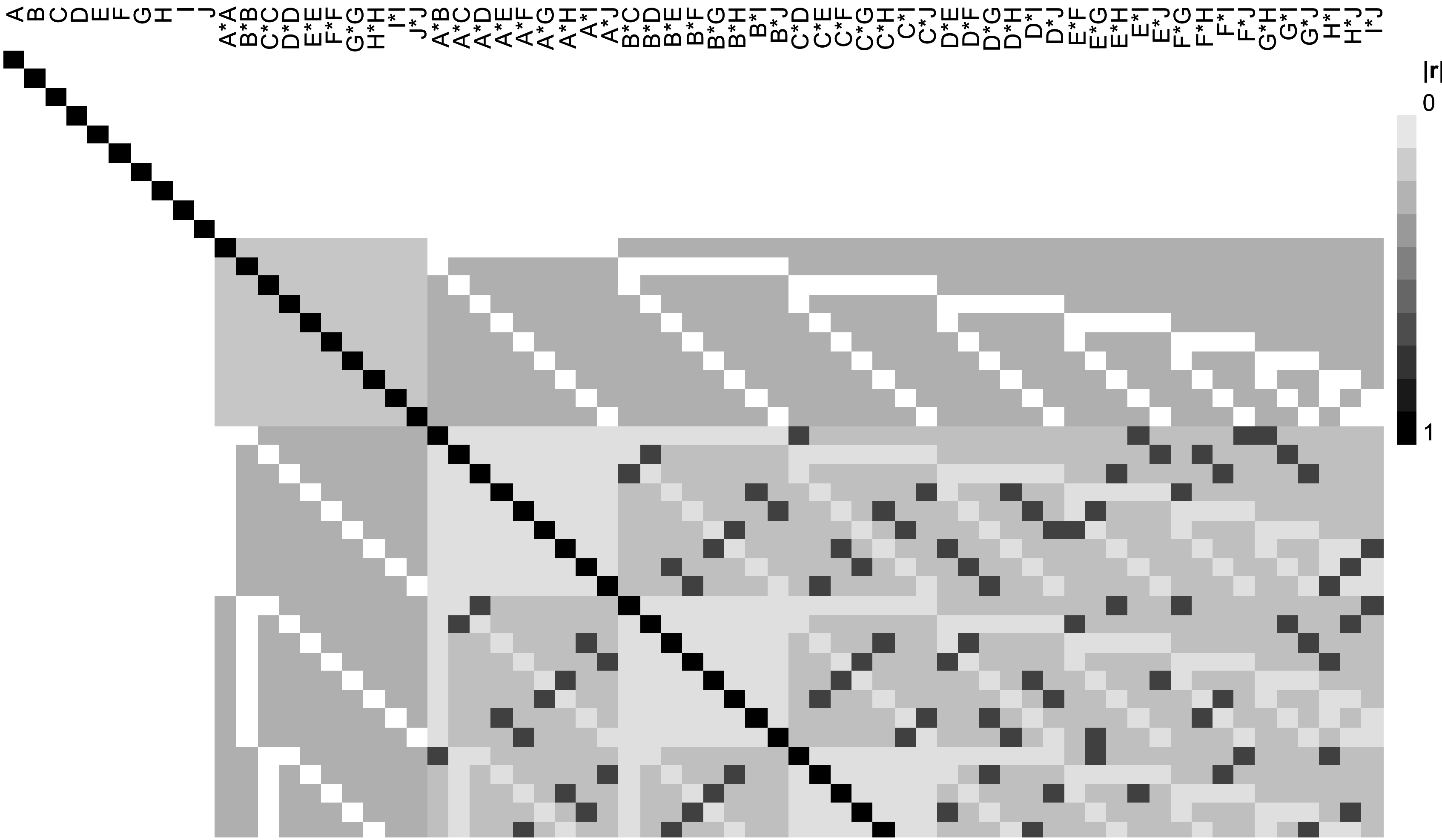
10-FACTOR DEFINITIVE SCREENING DESIGN, PROJECTION IN ALL 2-FACTOR COMBINATIONS (LEFT) AND PROJECTION IN FIRST THREE FACTORS (RIGHT)



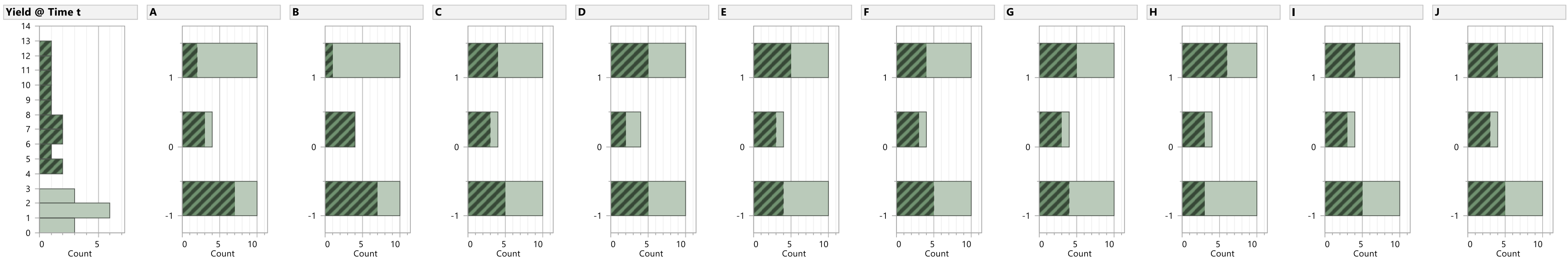


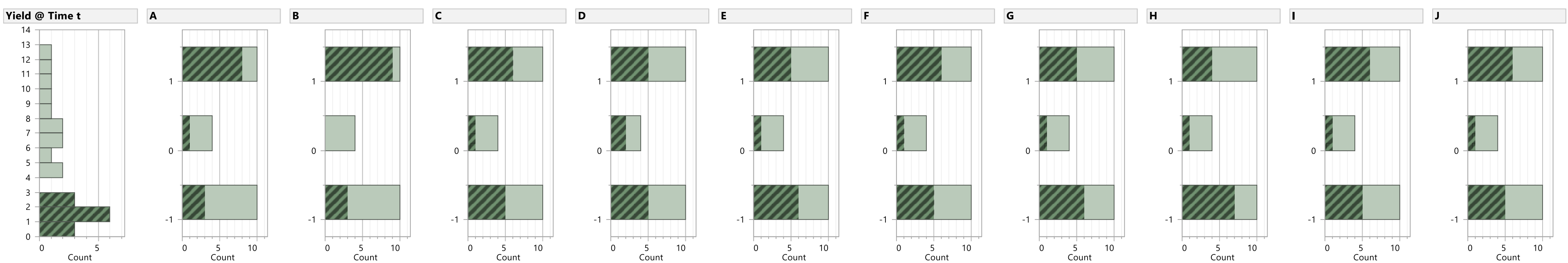
COLOR MAP FOR 10-FACTOR, 21-TRIAL, DEFINITIVE SCREENING DESIGN

Color Map On Correlations

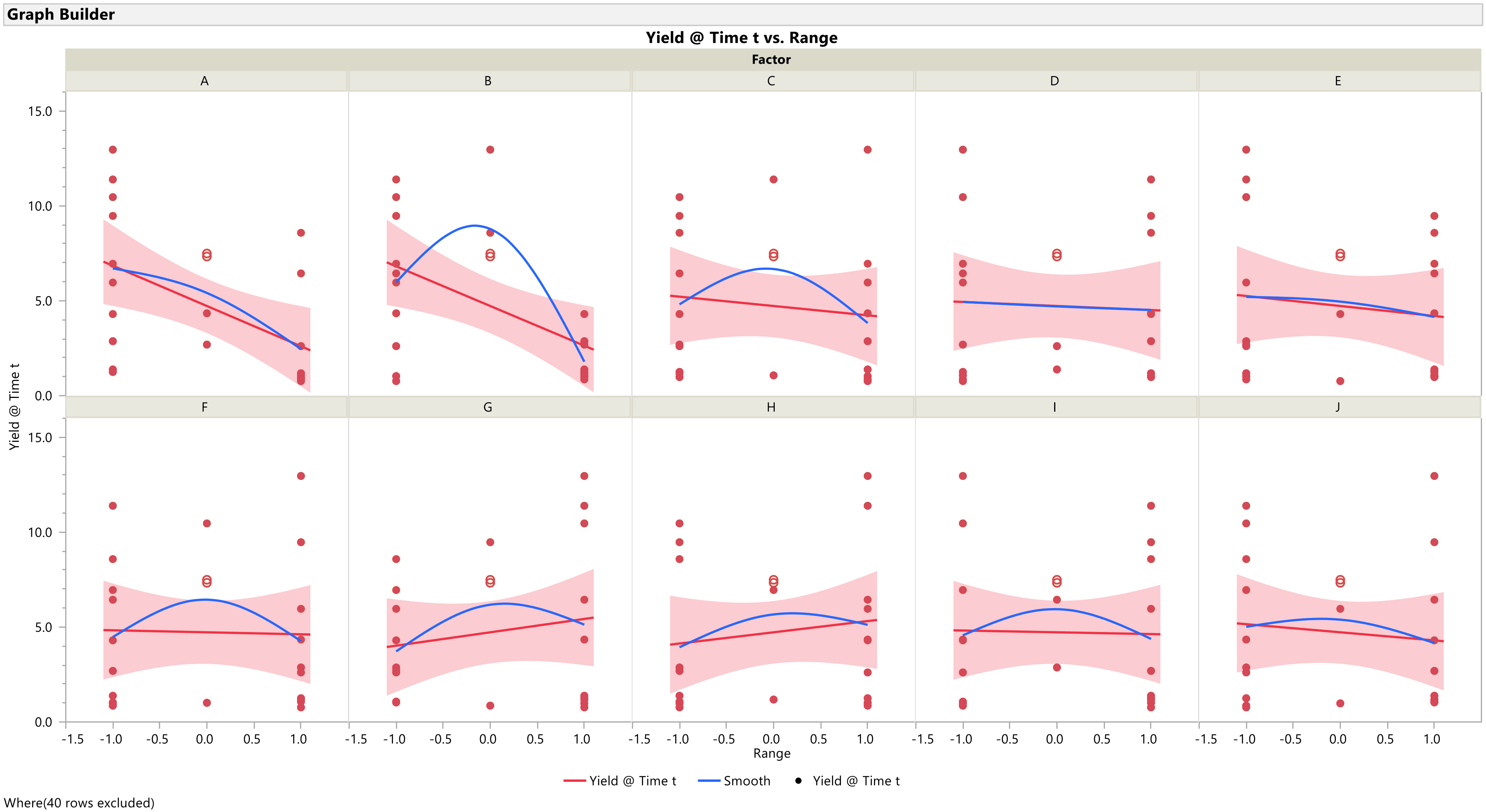


DISTRIBUTIONS WITH “GOOD” AND “BAD” BEHAVIOR SELECTED



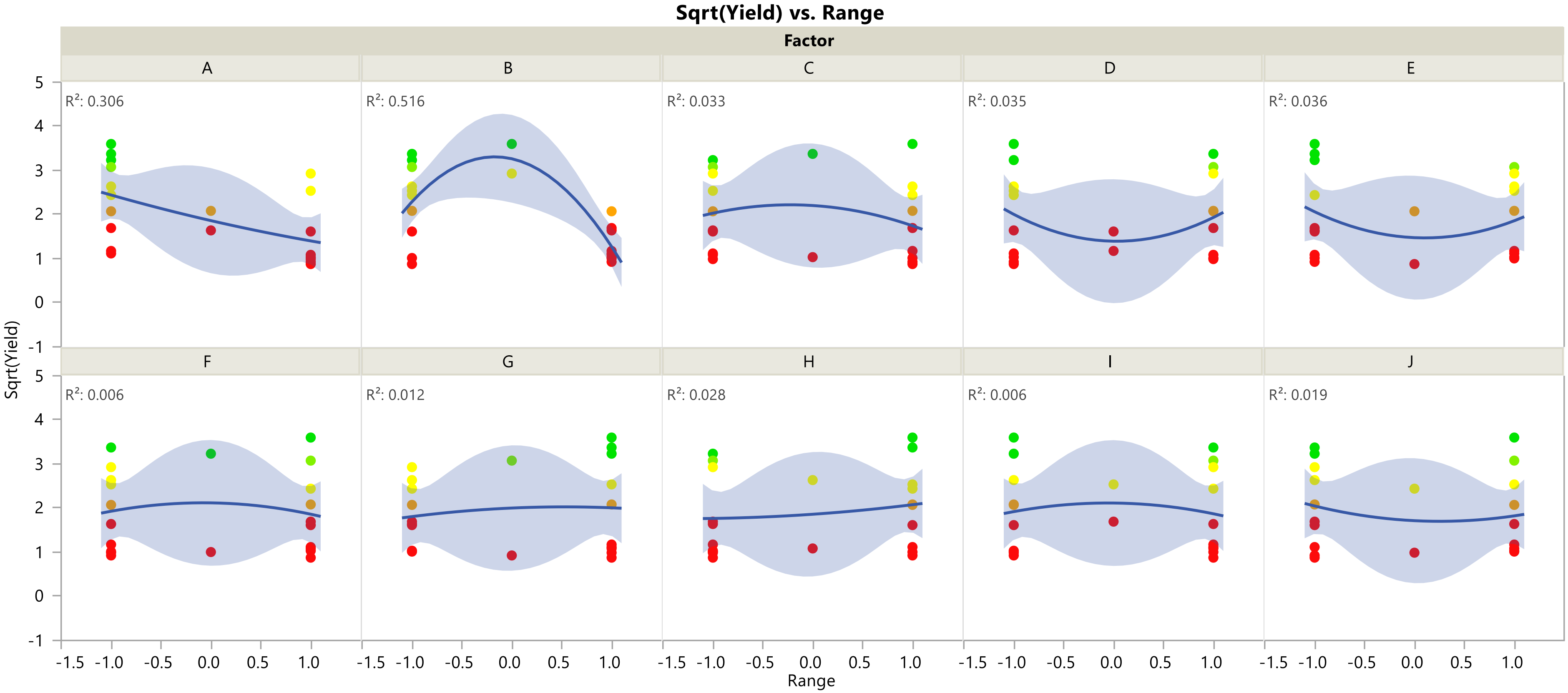


Y VS X PLOTS OF DATA FOR EACH X



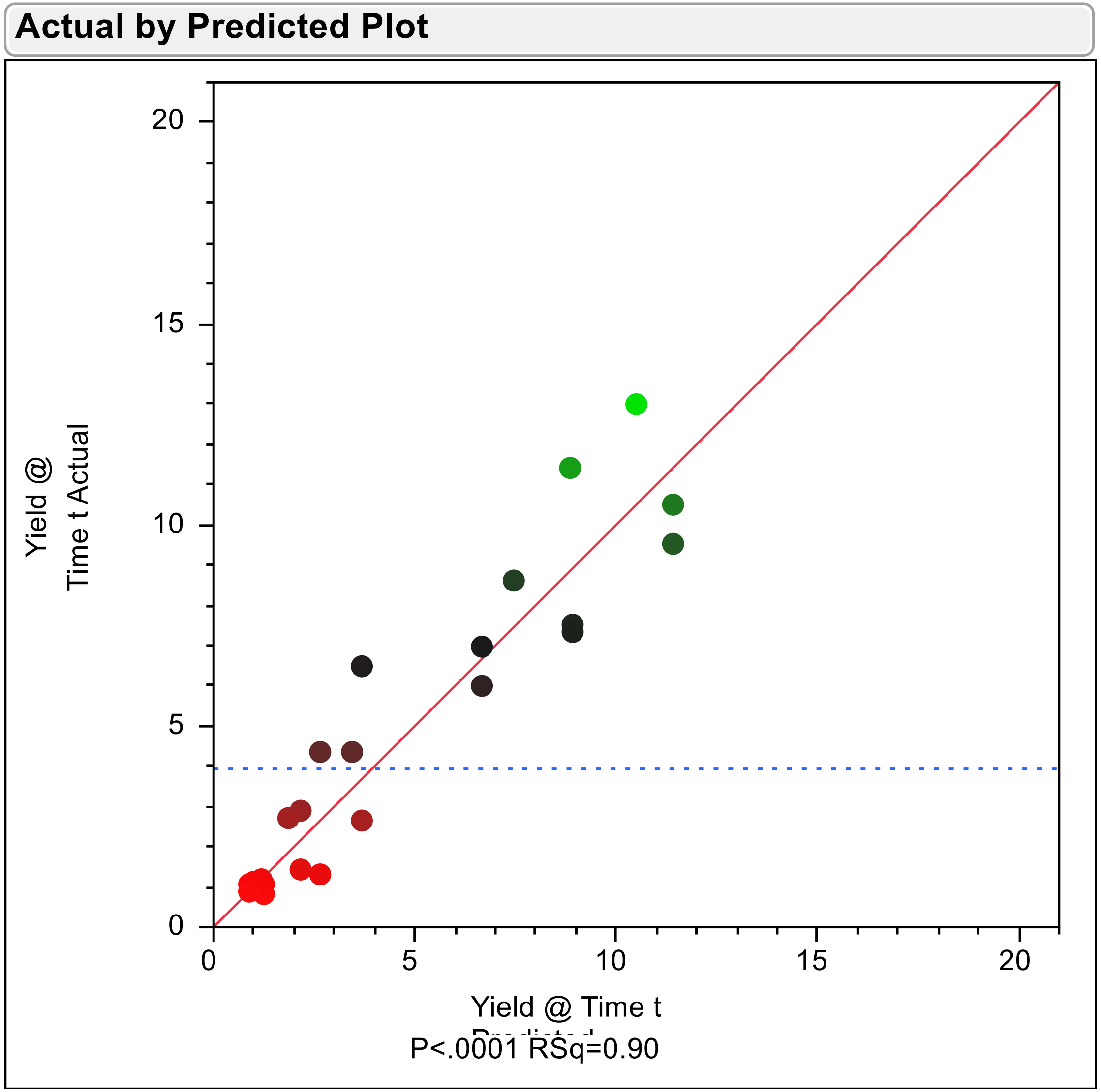
SQRT(Y) VS X PLOTS OF DATA FOR EACH X

Graph Builder

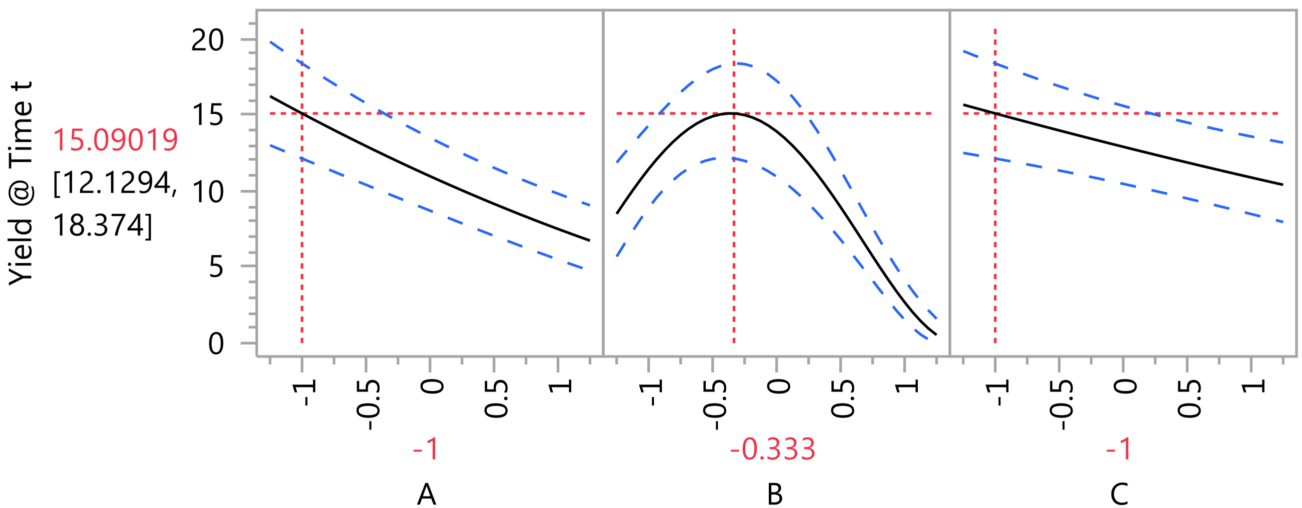


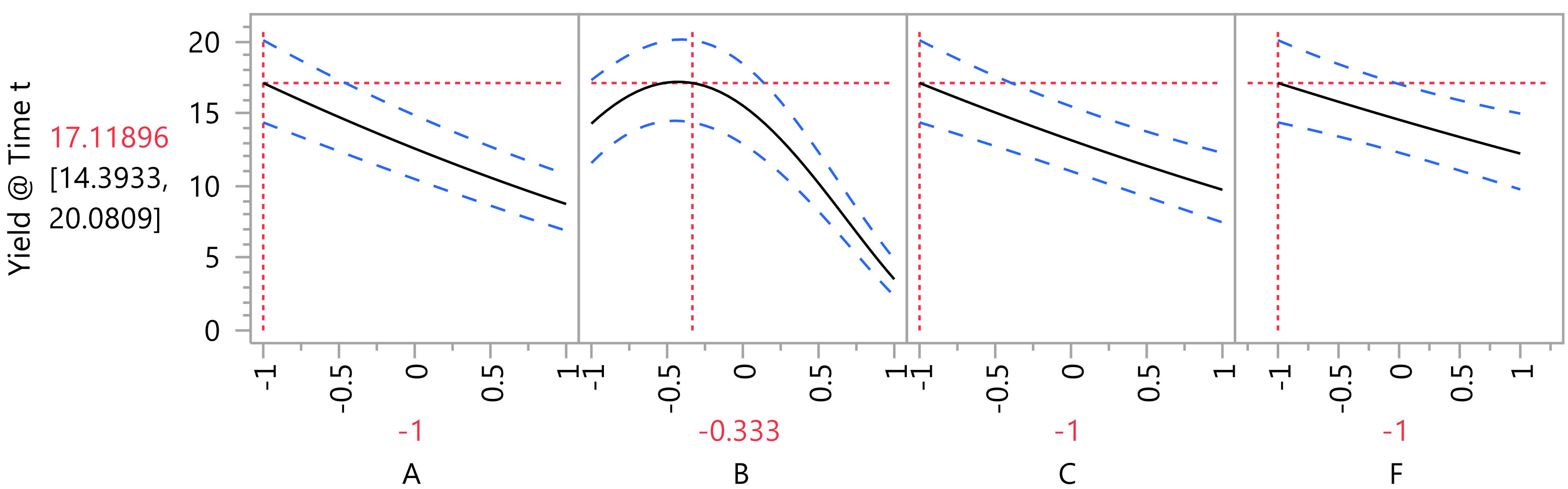
Where(60 rows excluded)

ACTUAL BY PREDICTED PLOT FOR FINAL 3-FACTOR MODEL FOR THE 24 DESIGN TRIALS

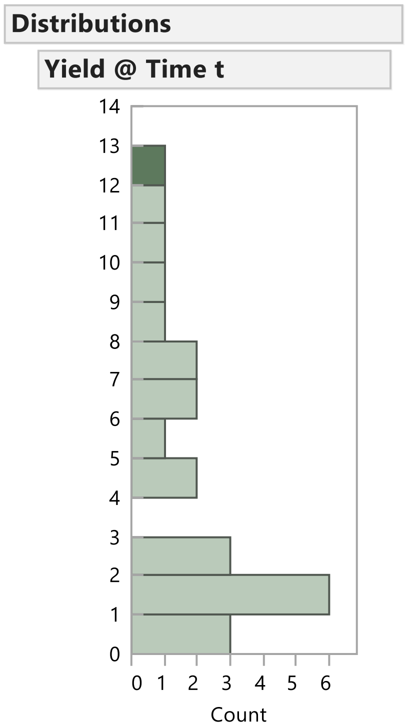


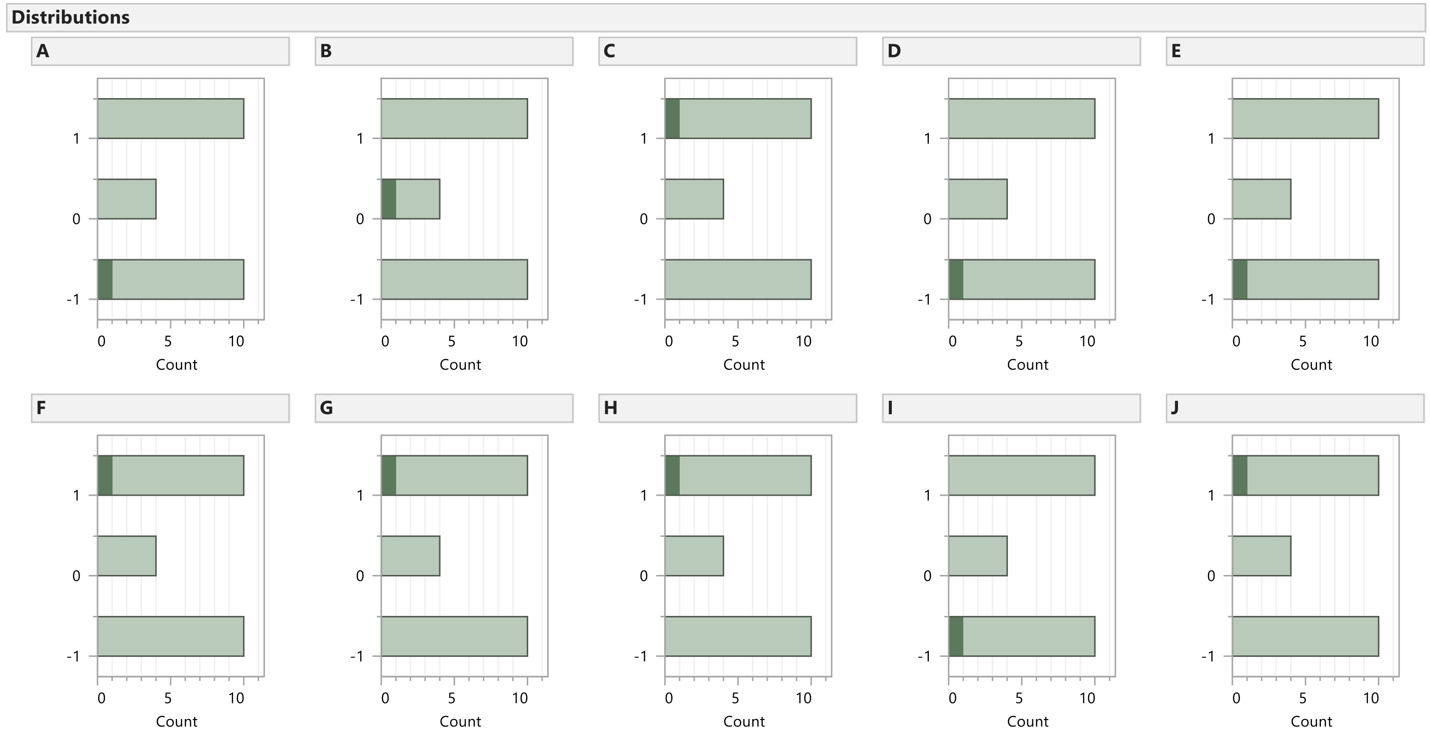
PREDICTING WITH BEST 3-FACTOR AND 4-FACTOR MODELS



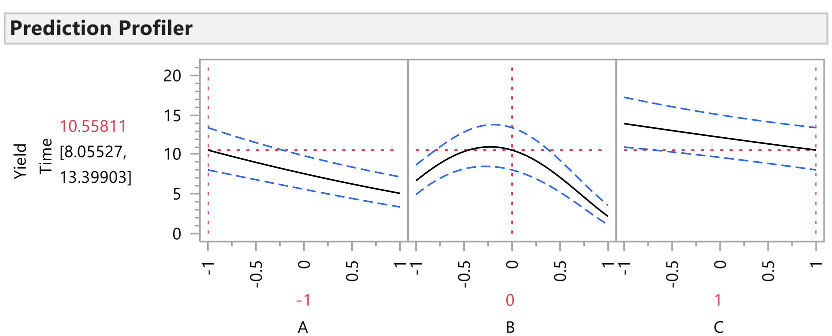


SETTINGS OF BEST OBSERVATION OF YIELD = 12.96

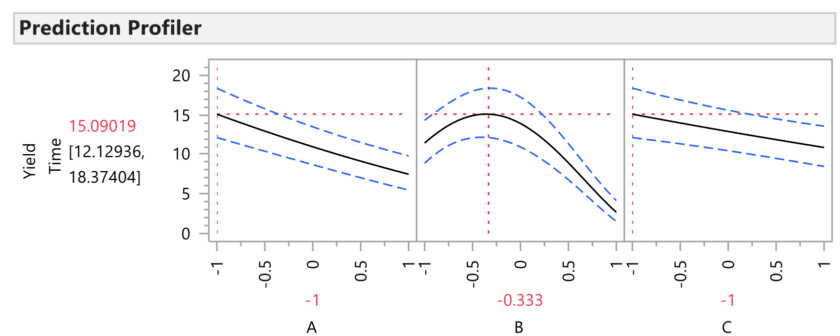




Prediction at settings of best observation

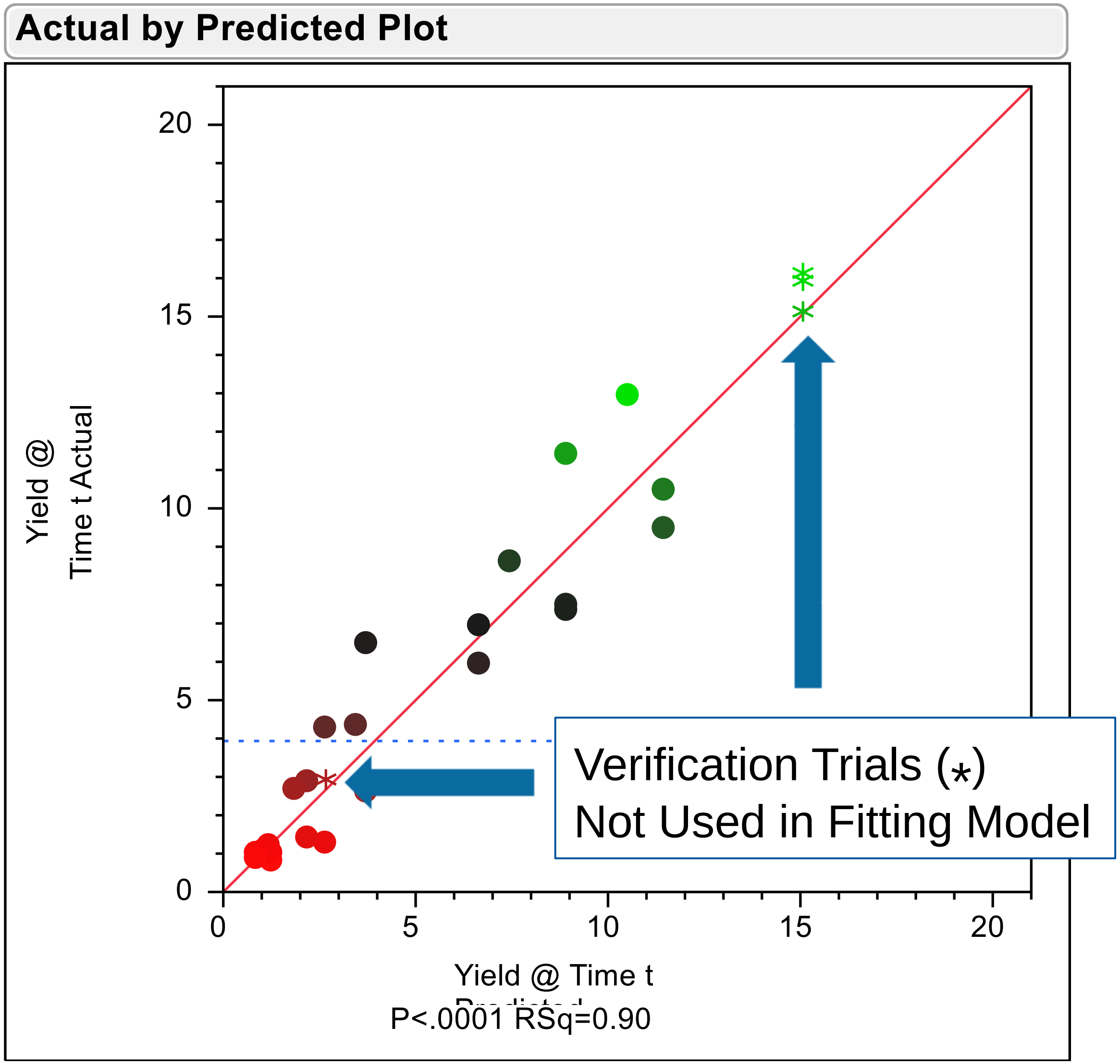


Prediction at best settings – run this checkpoint



23/1	Yield @ Time t	A	B	C	D	E	F	G	H	I	J
1	1.38	-1	1	1	0	1	-1	1	-1	1	1
2	6.44	1	-1	-1	-1	1	-1	1	1	0	1
3	5.96	-1	-1	1	-1	-1	1	-1	1	1	0
4	4.34	0	-1	1	1	1	1	1	1	-1	-1
5	10.46	-1	-1	-1	-1	-1	0	1	-1	-1	-1
6	6.95	-1	-1	1	-1	1	-1	-1	0	-1	-1
7	8.58	1	0	-1	1	1	-1	-1	-1	1	-1
8	2.69	0	1	-1	-1	-1	-1	-1	-1	1	1
9	4.3	-1	1	-1	1	0	-1	-1	1	-1	1
10	0.77	1	-1	1	-1	0	1	1	-1	1	-1
11	2.87	-1	1	1	1	-1	1	-1	-1	0	-1
12	1.01	1	1	1	1	1	0	-1	1	1	1
13	9.47	-1	-1	-1	1	1	1	0	-1	1	1
14	7.49	0	0	0	0	0	0	0	0	0	0
15	0.98	1	1	-1	1	1	-1	1	-1	-1	0
16	0.86	1	1	1	-1	-1	-1	0	1	-1	-1
17	1.25	-1	1	-1	-1	1	1	1	1	1	-1
18	1.03	1	-1	1	1	-1	-1	-1	-1	-1	1
19	1.07	1	1	0	-1	1	1	-1	-1	-1	1
20	7.33	0	0	0	0	0	0	0	0	0	0
21	2.61	1	-1	-1	0	-1	1	-1	1	-1	-1
22	11.39	-1	-1	0	1	-1	-1	1	1	1	-1
23	12.96	-1	0	1	-1	-1	1	1	1	-1	1
24	1.18	1	1	-1	1	-1	1	1	0	1	1
25	15.93	-1	-0.333	-1	1	-1	-1	1	1	1	1
26	2.9	-1	1	-1	1	-1	-1	1	1	1	1
27	16.16	-1	-0.333	-1	-1	-1	-1	1	1	1	1
28	15.1	-1	-0.333	-1	0	-1	-1	1	1	1	1

ACTUAL BY PREDICTED PLOT FOR FINAL 3-FACTOR MODEL FOR THE 24 DESIGN TRIALS AND 4 VERIFICATION TRIALS



DISCOVERY SUMMIT VIDEO INTRO TO FIT DEFINITIVE SCREENING

- 2017 JMP Discovery Summit presentation by Brad Jones on
  - Simulating Responses and Fitting Definitive Screening Designs - JMP User Community

MAIN EFFECT RANKING OF FACTORS

NEW DEFINITIVE SCREENING ANALYSIS METHOD

Effect Summary

Source	LogWorth	PValue
A	1.622	0.02387
B	1.568	0.02705
C	0.515	0.30541
F	0.239	0.57657
J	0.169	0.67802
G	0.159	0.69271
H	0.141	0.72196
E	0.141	0.72231
I	0.071	0.84905
D	0.061	0.86836

- Treat factors D and I as the dummy factors to be used for error estimates in Definitive Screening Fit

DSD FIT OUTPUT WITH FACTORS D & I USED FOR ERROR

NEW DEFINITIVE SCREENING ANALYSIS METHOD

Stage 1 - Main Effect Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
A	-2.05	0.2228	-9.2	<.0001*
B	-2.015	0.2228	-9.043	<.0001*
C	-0.855	0.2228	-3.839	0.0050*
F	-0.427	0.2228	-1.916	0.0917

Statistic	Value
RMSE	0.9839
DF	8

Stage 2 - Even Order Effect Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	8.6319	0.6421	13.442	<.0001*
A*B	1.2645	0.2968	4.2608	0.0037*
B*C	0.9481	0.3036	3.1232	0.0168*
B*F	0.5687	0.3036	1.8733	0.1032
C*F	0.9163	0.3213	2.8517	0.0246*
B*B	-4.756	0.7043	-6.753	0.0003*

Statistic	Value
RMSE	1.2435
DF	7

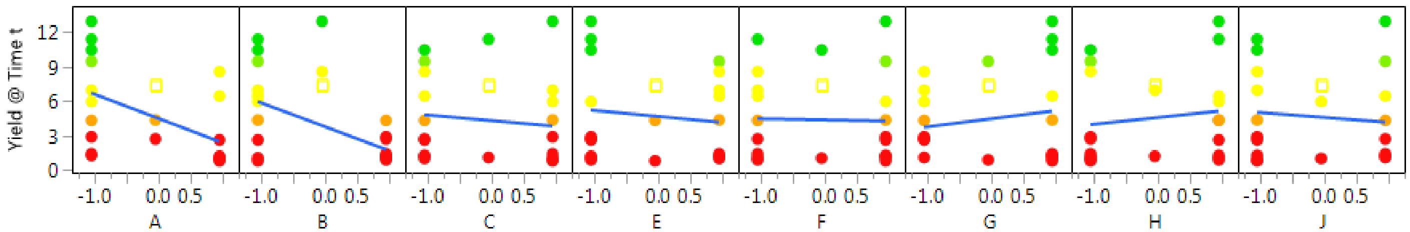
Combined Model Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	8.6319	0.5947	14.514	<.0001*
A	-2.05	0.2608	-7.86	<.0001*
B	-2.015	0.2608	-7.726	<.0001*
C	-0.855	0.2608	-3.279	0.0055*
F	-0.427	0.2608	-1.637	0.1239
A*B	1.2645	0.2749	4.6006	0.0004*
B*C	0.9481	0.2812	3.3722	0.0046*
B*F	0.5687	0.2812	2.0227	0.0626
C*F	0.9163	0.2976	3.0791	0.0082*
B*B	-4.756	0.6523	-7.292	<.0001*

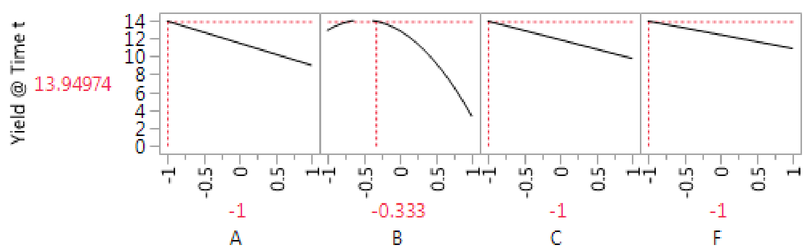
Statistic	Value
RMSE	1.1516
DF	14

Make Model Run Model

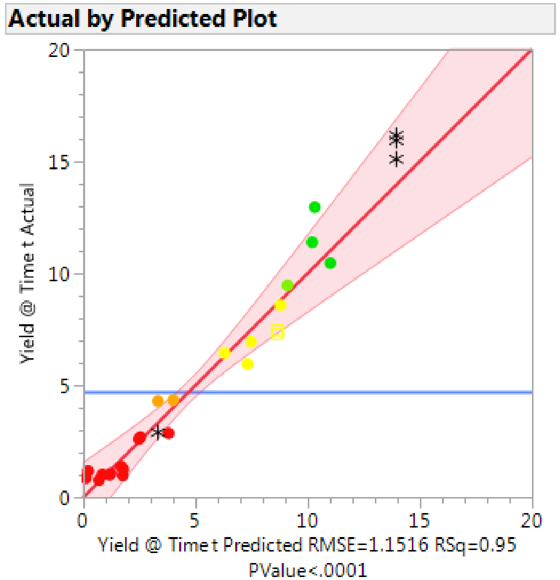
Main Effects Plot

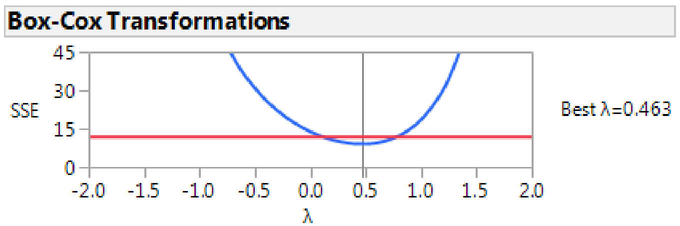


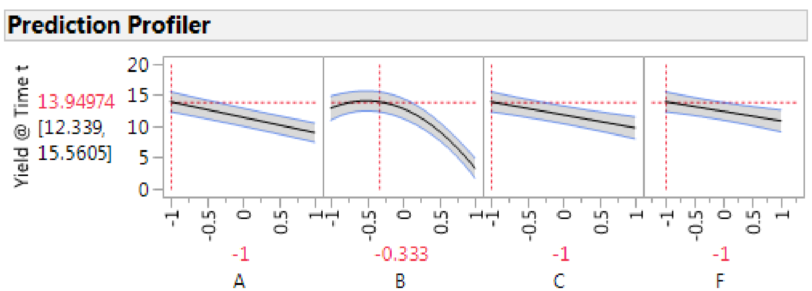
Prediction Profiler



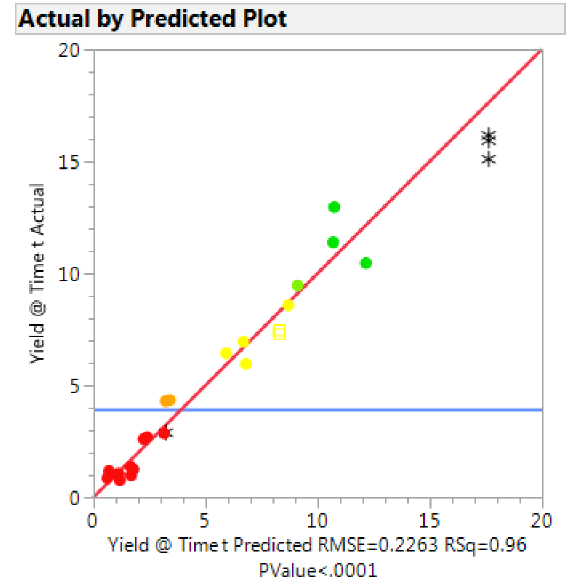
FIT OF RAW YIELD

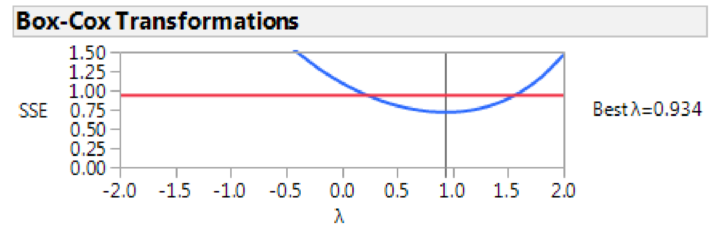


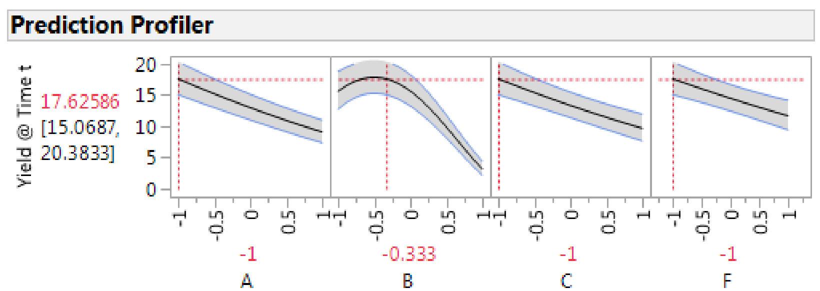




FIT OF SQRT YIELD







ANALYSIS STRATEGIES FOR WHEN YOU DON'T HAVE THE NEW DEFINITIVE SCREENING ANALYSIS METHOD

- Conservative – start by treating designs like traditional screening
  - Fit main effects only – DSD is orthogonal in main effects
  - Then fit ME + squared effects – DSD is orthogonal in squared terms too
  - *Use *factor sparsity* and *effect heredity* principles to propose final models
  - Use transformation to make error more uniform
    - » square-root identified in plot of SSE vs. λ for Box-Cox transformation (i.e. λ ≈ 0.5)

- Aggressive – use stepwise regression to pick "best" subsets of terms
  - Use AICc & BIC stopping criteria and pick "simpler model" – Occam's razor
  - Use max K-Fold R-square as stopping rule to pick model (no checkpoints)
  - Use max validation R-square for checkpoints as stopping rule to pick model
  - Fit ALL possible models

**Factor sparsity* states only a few variables will be active in a factorial DOE
Effect heredity states significant interactions will only occur if at least one parent is active
Pg. 112 , Wu & Hamada, "*Experiments, Planning, Analysis and Parameter Design Optimization*"

ALL ANALYSES RANK FACTORS A, B & C AS TOP 3

FACTOR F APPEARS TO BE MOST LIKELY FOURTH FACTOR

- Linear terms only – fourth factor is F
- Linear + Squared terms – fourth factor is D
- Stepwise with min AICc stopping rule – fourth factor is F
- Stepwise with max K-Fold R-Square stopping rule – fourth factor is F
- Stepwise with max Validation R-Square as stopping rule – fourth factor is F
- All possible models – fourth factor is G

- When D & F are in same 5-factor (with A, B, & C) stepwise model, D drops out
- When G & F are in same 5-factor (with A, B, & C) stepwise model, G drops out
- When D & G are in same 5-factor (with A, B, & C) stepwise model, both drop out

- There is an important difference between saying, *"Factor F has no effect."* and, *"Given the amount of data taken an effect for factor F was not detected."*

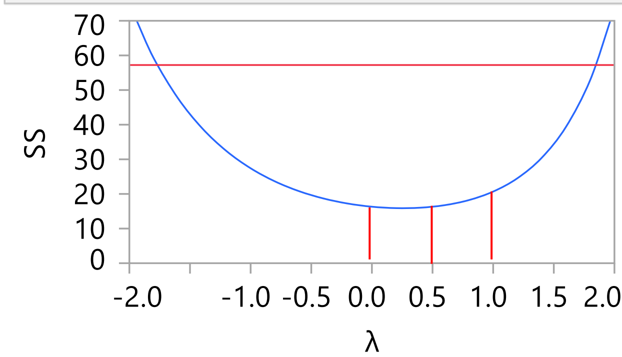
- Augmenting design to support 6-factor quadratic model in A, B, C, D, F & G will
  - help resolve the relative contributions of D, F & G
  - increase the power for all – but especially - the squared terms

CONSERVATIVE ANALYSIS

Sorted Parameter Estimates

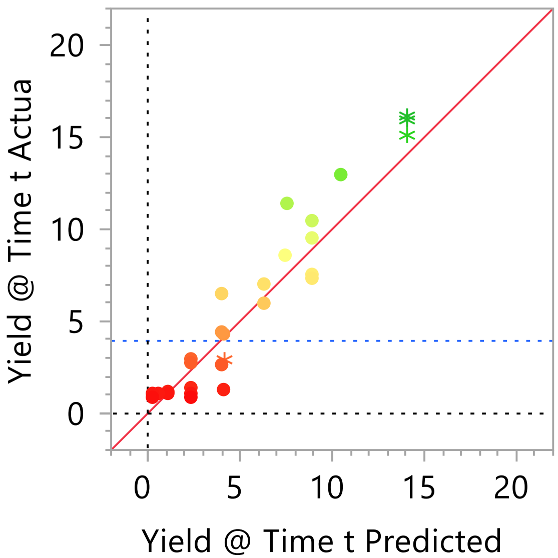
Term	Estimate	Std Error	t Ratio	Prob>|t|
A	-2.023428	0.791305	-2.56	0.0239 *
B	-2.030884	0.815352	-2.49	0.0271 *
C	-0.844283	0.791305	-1.07	0.3054
F	-0.453239	0.791305	-0.57	0.5766
J	0.3462584	0.815352	0.42	0.6780
G	0.3230058	0.799335	0.40	0.6927
H	0.2867159	0.788411	0.36	0.7220
E	-0.287384	0.791305	-0.36	0.7223
I	-0.155204	0.799335	-0.19	0.8490
D	0.1332841	0.788411	0.17	0.8684

Box-Cox Transformations



TRANSFORMATIONS
SQRT, LOG, & NONE

Actual by Predicted Plot

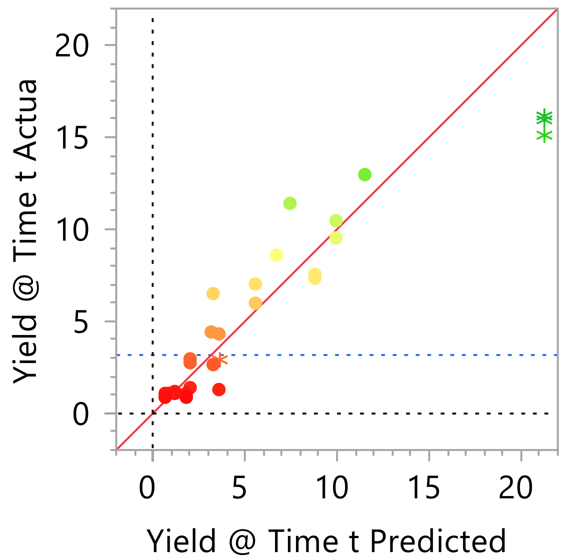


P<.0001 RSq=0.83
RMSE=0.4163

Summary of Fit

RSquare	0.825967
RSquare Adj	0.789328
Root Mean Square Error	0.416337
Mean of Response	1.983747
Observations (or Sum Wgts	24

Actual by Predicted Plot

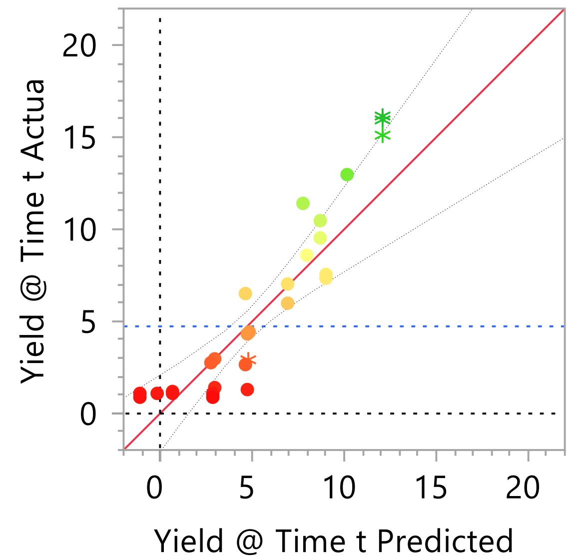


P<.0001 RSq=0.82
RMSE=0.4509

Summary of Fit

RSquare	0.823029
RSquare Adj	0.785772
Root Mean Square Error	0.450888
Mean of Response	1.151951
Observations (or Sum Wgts	24

Actual by Predicted Plot



P<.0001 RSq=0.79
RMSE=1.9387

Summary of Fit

RSquare	0.789957
RSquare Adj	0.745738
Root Mean Square Error	1.938688
Mean of Response	4.72375
Observations (or Sum Wgts	24

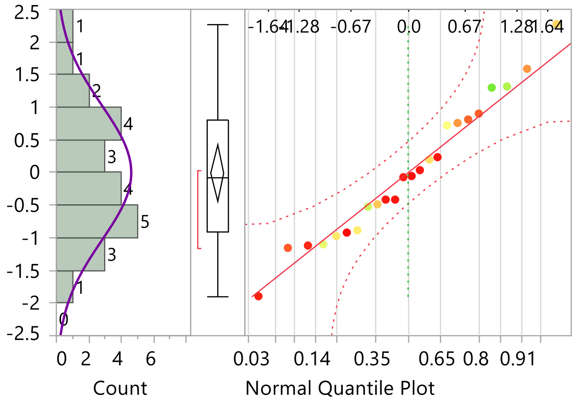
PLOTS OF RESIDUALS FOR DIFFERENT TRANSFORMATIONS

Model fit was reduced quadratic in A, B & C:

Yield = Intercept + A + B + C + B*B + A*B + B*C

Distributions

Studentized Resid Sqrt(Yield @ Time t) 2



Normal(-0.0045,1.03596)

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	-0.004478	-0.441926	0.4329688
Dispersion	σ	1.0359592	0.8051616	1.4532028

-2log(Likelihood) = 68.8047829349136

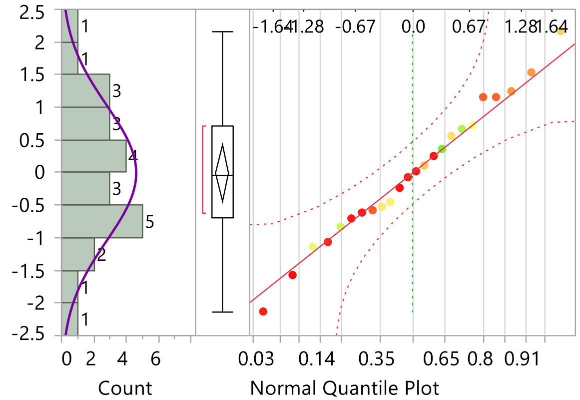
Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.972241	0.7224

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Studentized Resid Log(Yield @ Time t)



Normal(-0.008,1.03586)

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	-0.007981	-0.445387	0.4294258
Dispersion	σ	1.035863	0.8050868	1.4530679

-2log(Likelihood) = 68.8003267780461

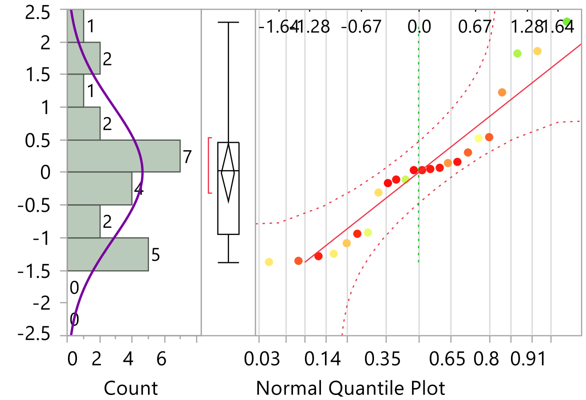
Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.992406	0.9994

Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

Studentized Resid Yield @ Time t



Normal(-0.0003,1.0284)

Fitted Normal

Parameter Estimates

Type	Parameter	Estimate	Lower 95%	Upper 95%
Location	μ	-0.000276	-0.434534	0.4339807
Dispersion	σ	1.0284046	0.79929	1.4426054

-2log(Likelihood) = 68.4534641248215

Goodness-of-Fit Test

Shapiro-Wilk W Test

W	Prob<W
0.918997	0.0555

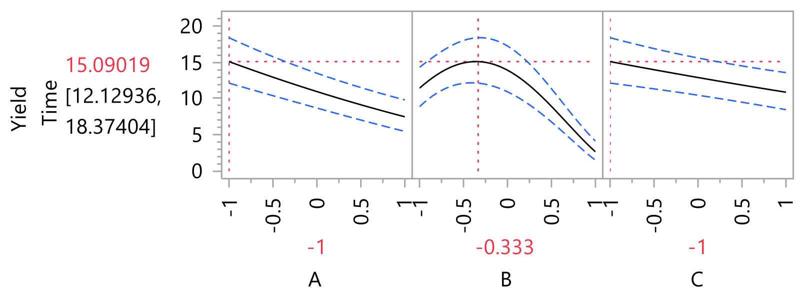
Note: Ho = The data is from the Normal distribution. Small p-values reject Ho.

STEPWISE 3-FACTOR MODEL (7 TERMS) - LEFT
FULL QUADRATIC 3-FACTOR MODEL (10 TERMS) - RIGHT

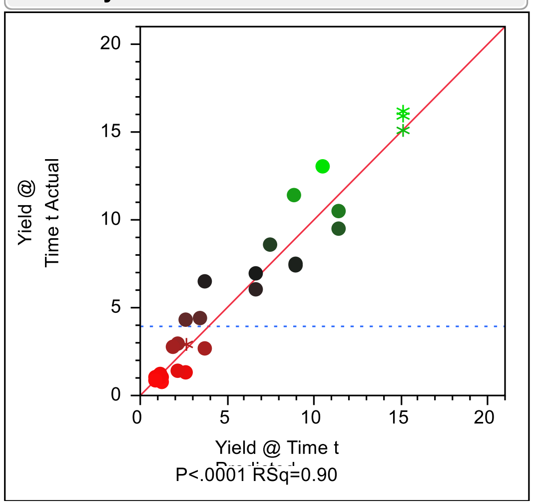
Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
B*B	-1.218717	0.182702	-6.67	<.0001 *
A	-0.496169	0.075133	-6.60	<.0001 *
B	-0.481867	0.075133	-6.41	<.0001 *
C	-0.240181	0.075133	-3.20	0.0053 *
A*B	0.2306449	0.078918	2.92	0.0095 *
C*B	0.1585526	0.078918	2.01	0.0607

Prediction Profiler



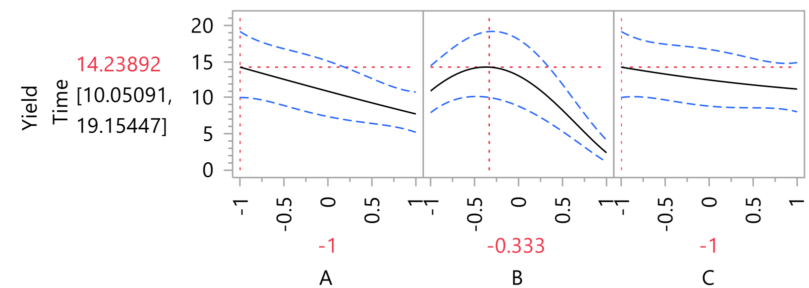
Actual by Predicted Plot



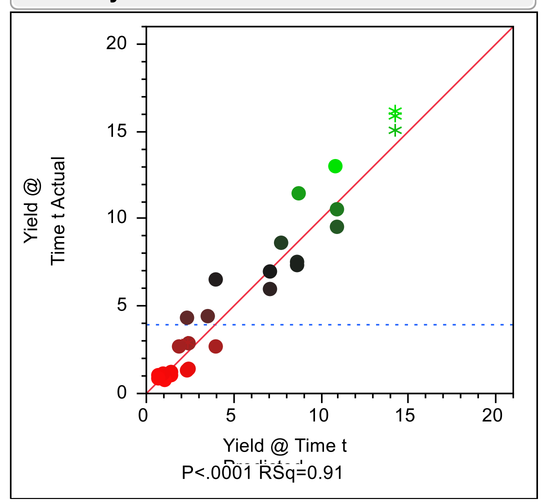
Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
A	-0.496169	0.080197	-6.19	<.0001 *
B	-0.481867	0.080197	-6.01	<.0001 *
B*B	-1.181941	0.233332	-5.07	0.0002 *
C	-0.240181	0.080197	-2.99	0.0096 *
A*B	0.2339616	0.087698	2.67	0.0184 *
C*B	0.1610152	0.087698	1.84	0.0877
A*C	-0.08124	0.087698	-0.93	0.3700
C*C	0.0307046	0.233332	0.13	0.8972
A*A	-0.021309	0.233332	-0.09	0.9285

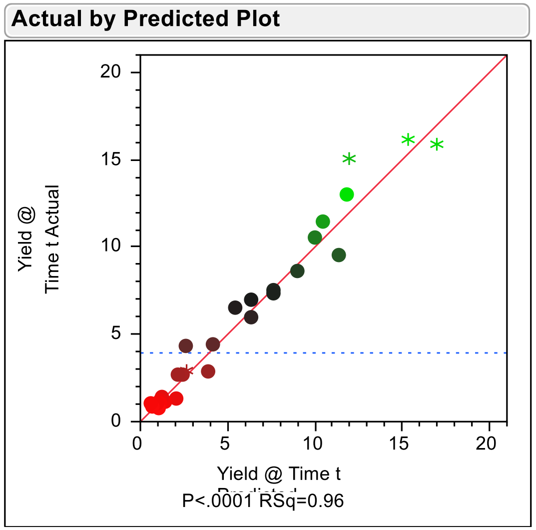
Prediction Profiler

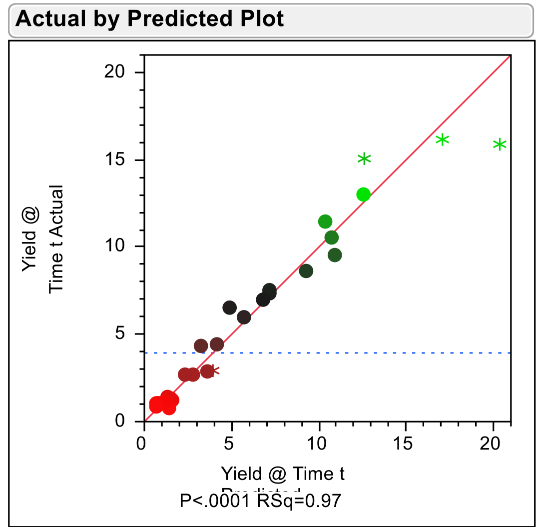


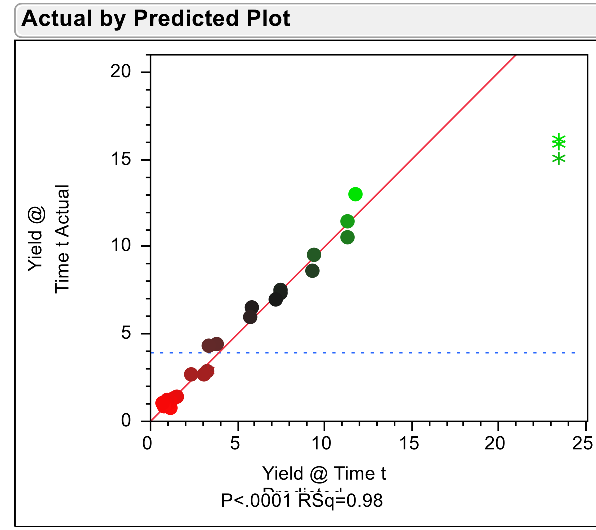
Actual by Predicted Plot

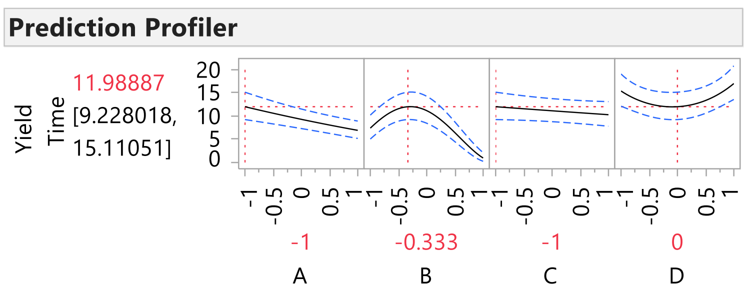


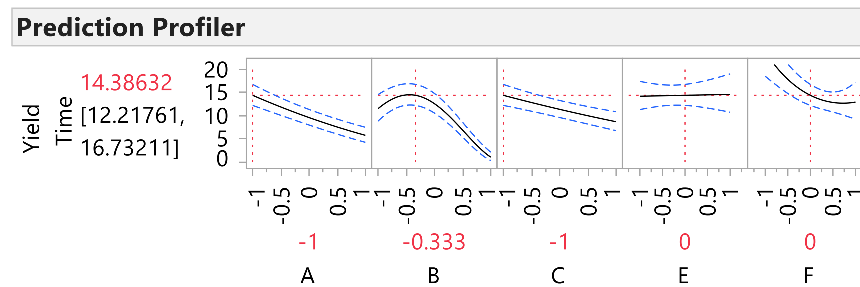
STEPWISE MODELS:
4-FACTOR (12 TERMS), 5-FACTOR (13 TERMS), 6-FACTOR (15 TERMS)

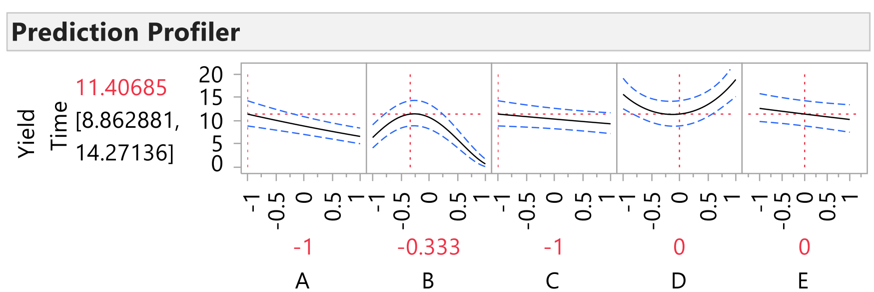












AGGRESSIVE ANALYSES

- Stepwise using Main Effects and Squared Effects for all factors
  - Will show just the use of AICc & BIC stopping criteria – all stepwise approaches yield very similar results

- Stepwise using full 10-factor, 66-term quadratic model

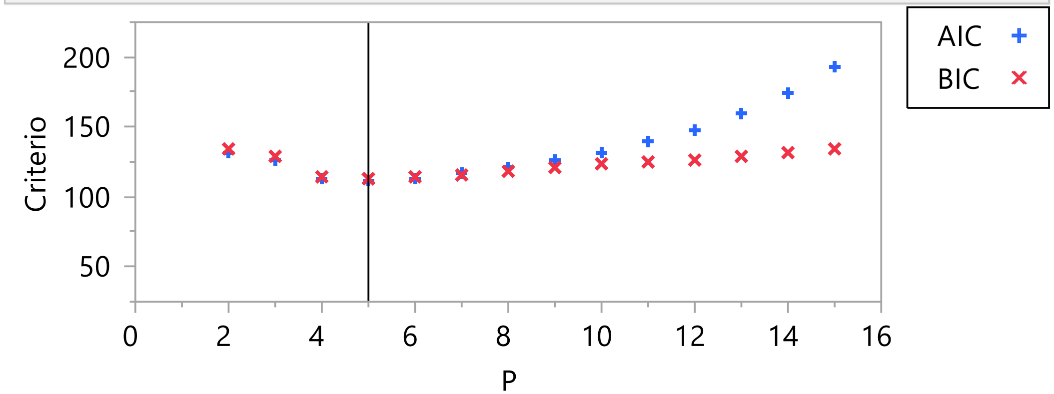
  1 intercept + 10 ME + 10 SQ + 45 2FI (2-factor interactions)
  - Use AICc & BIC stopping criteria and pick “simpler model” – Occam’s razor
  - Use max K-Fold R-square as stopping rule to pick model (no checkpoints)
  - Use max validation R-square for checkpoints as stopping rule to pick model
  - Fit ALL possible models

USE MIN AIC OR BIC CRITERION AS STOPPING RULE

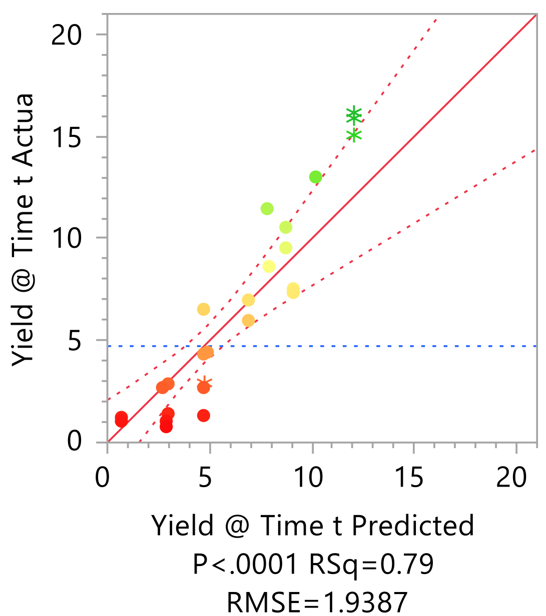
21 TERMS, ME + SQ

RAW RESPONSE VALUES USED

Criterion History



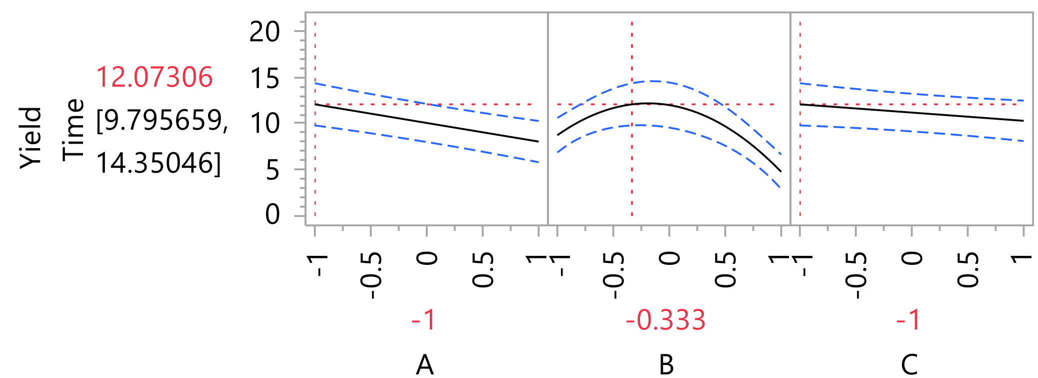
Actual by Predicted Plot



Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
B*B	-5.2395	1.061863	-4.93	<.0001 *
A	-2.014167	0.437499	-4.60	0.0002 *
B	-1.979167	0.437499	-4.52	0.0002 *
C	-0.890833	0.437499	-2.04	0.0559

Prediction Profiler

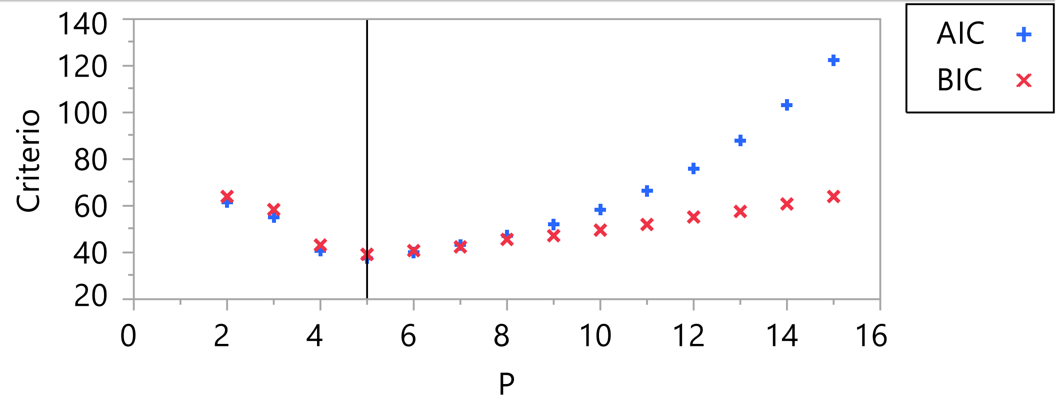


USE MIN AIC OR BIC CRITERION AS STOPPING RULE

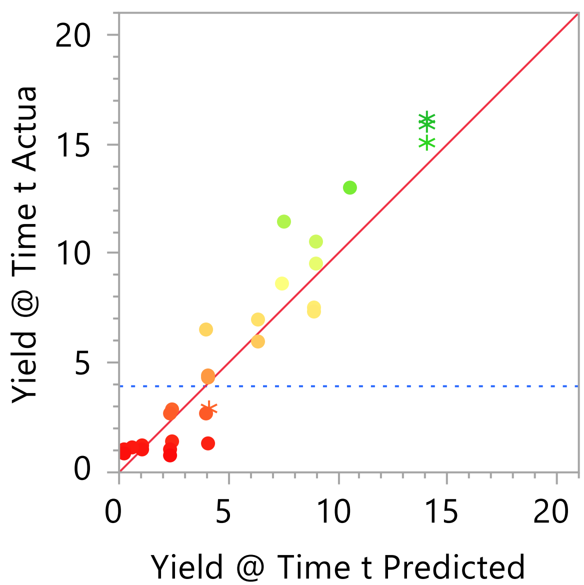
21 TERMS, ME + SQ

TRANSFORMED VALUES USED

Criterion History



Actual by Predicted Plot

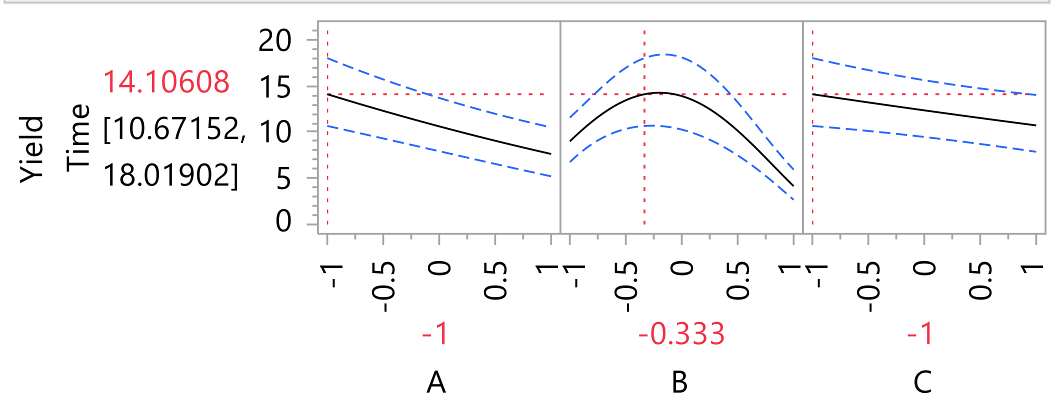


P<.0001 RSq=0.83
RMSE=0.4163

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
B*B	-1.211508	0.228037	-5.31	<.0001 *
A	-0.496169	0.093954	-5.28	<.0001 *
B	-0.481867	0.093954	-5.13	<.0001 *
C	-0.240181	0.093954	-2.56	0.0193 *

Prediction Profiler

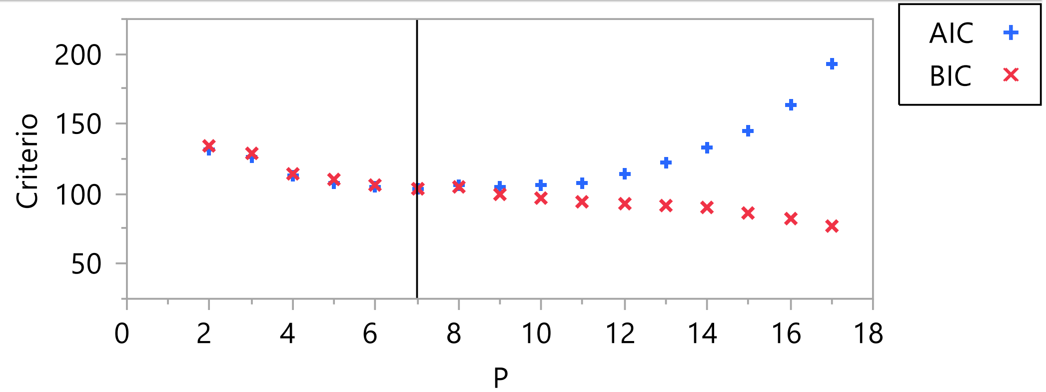


USE MIN AIC OR BIC CRITERION AS STOPPING RULE

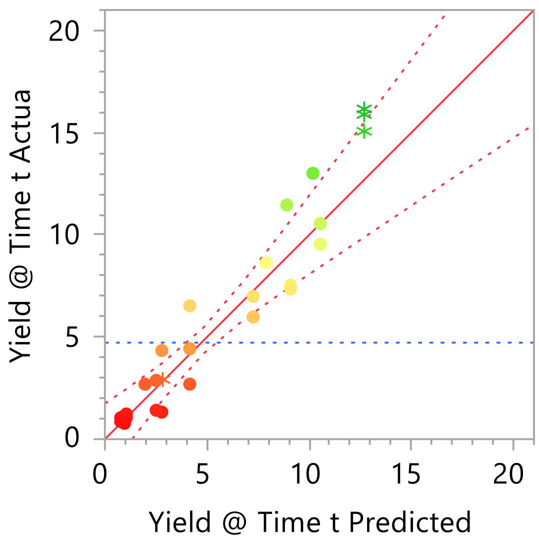
66 TERM QUADRATIC

RAW RESPONSE VALUES USED

Criterion History



Actual by Predicted Plot

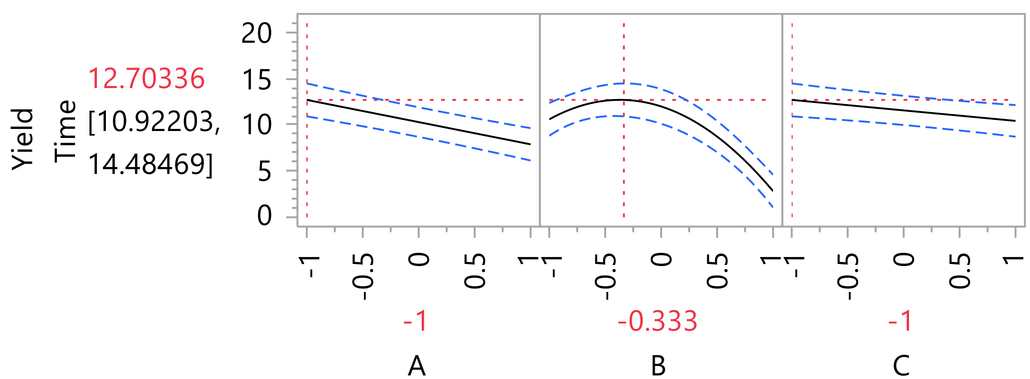


P<.0001 RSq=0.89
RMSE=1.4757

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
B*B	-5.282841	0.809809	-6.52	<.0001 *
A	-2.014167	0.33302	-6.05	<.0001 *
B	-1.979167	0.33302	-5.94	<.0001 *
A*B	1.1703157	0.349799	3.35	0.0038 *
C	-0.890833	0.33302	-2.68	0.0160 *
B*C	0.7369066	0.349799	2.11	0.0503

Prediction Profiler

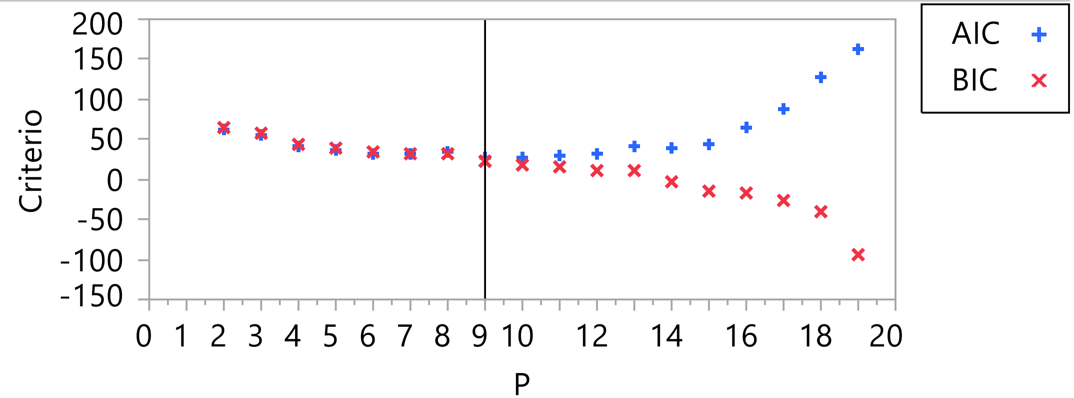


USE MIN AIC OR BIC CRITERION AS STOPPING RULE

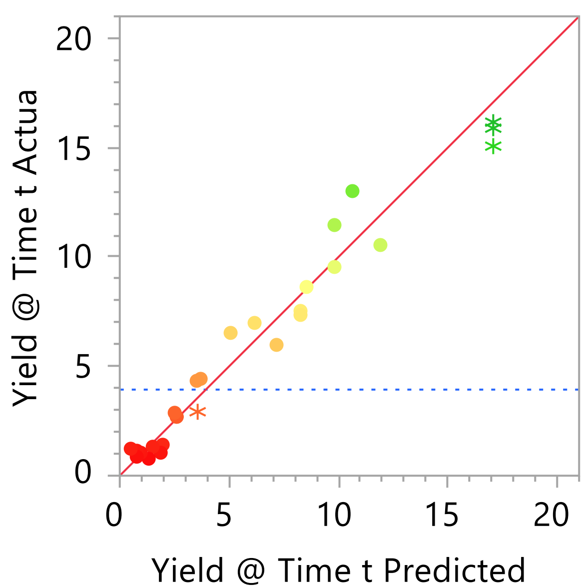
66 TERM QUADRATIC

TRANSFORMED VALUES USED

Criterion History



Actual by Predicted Plot

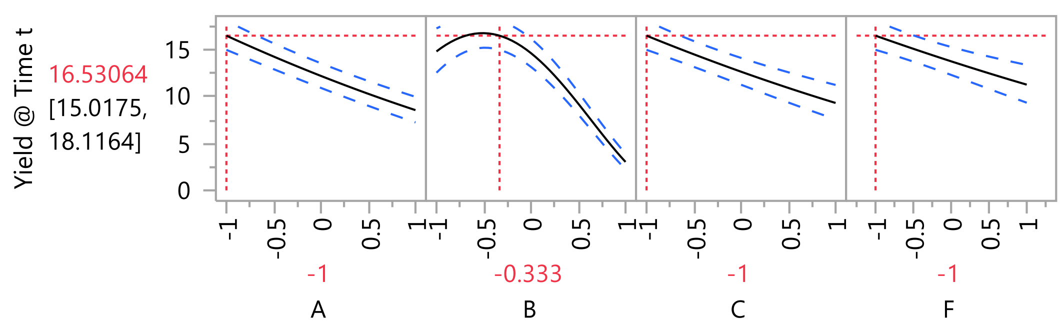


P<.0001 RSq=0.95
RMSE=0.2519

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
A	-0.505343	0.057053	-8.86	<.0001 *
B	-0.491041	0.057053	-8.61	<.0001 *
B*B	-1.111685	0.141981	-7.83	<.0001 *
A*B	0.253637	0.060121	4.22	0.0007 *
C	-0.231007	0.057053	-4.05	0.0010 *
B*C	0.2053297	0.061367	3.35	0.0044 *
C*F	0.2093075	0.063209	3.31	0.0047 *
F	-0.110087	0.057053	-1.93	0.0728

Prediction Profiler

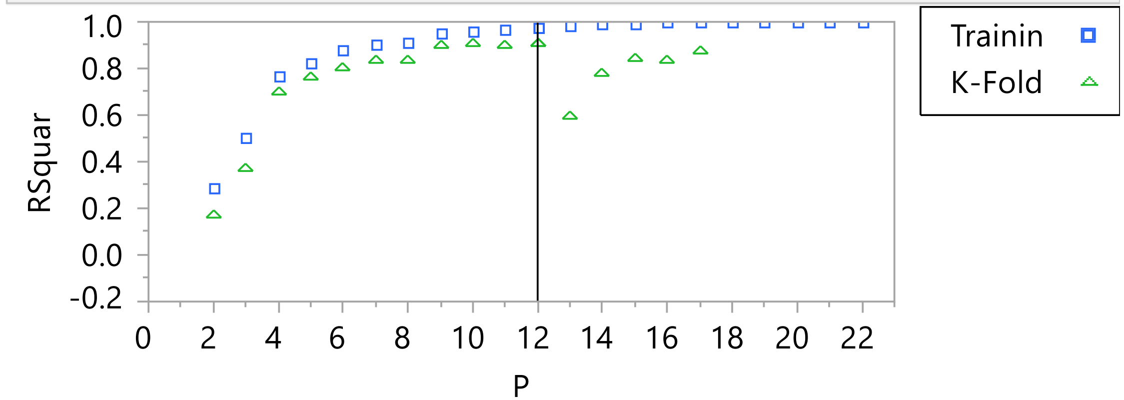


USE MAX K-FOLD R-SQUARE AS STOPPING RULE

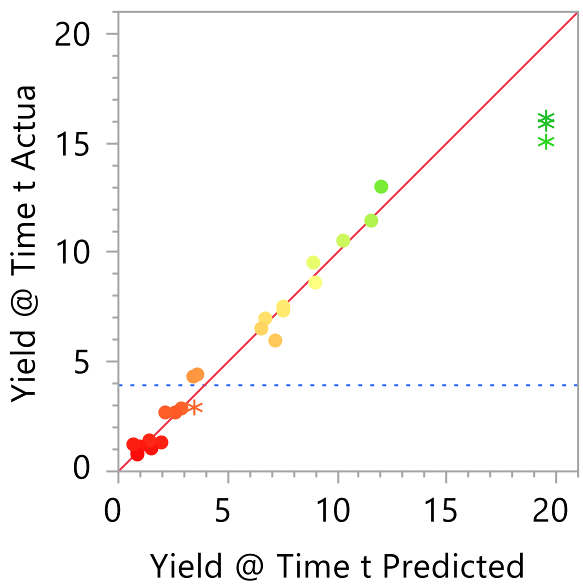
66 TERM QUADRATIC

TRANSFORMED VALUES USED

RSquare History



Actual by Predicted Plot

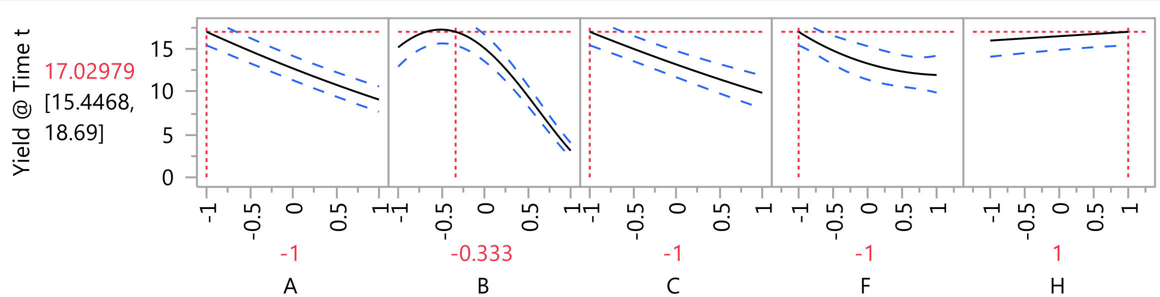


P<.0001 RSq=0.98
RMSE=0.1839

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
A	-0.498201	0.041762	-11.93	<.0001 *
B	-0.483899	0.041762	-11.59	<.0001 *
B*B	-1.184839	0.114991	-10.30	<.0001 *
A*B	0.2798015	0.045426	6.16	<.0001 *
C	-0.238149	0.041762	-5.70	<.0001 *
B*C	0.2427713	0.047097	5.15	0.0002 *
C*F	0.2349251	0.047559	4.94	0.0003 *
F	-0.117229	0.041762	-2.81	0.0158 *
B*F	0.1203014	0.0449	2.68	0.0201 *
H	0.0928467	0.041762	2.22	0.0462 *
F*F	0.2478009	0.120097	2.06	0.0614

Prediction Profiler

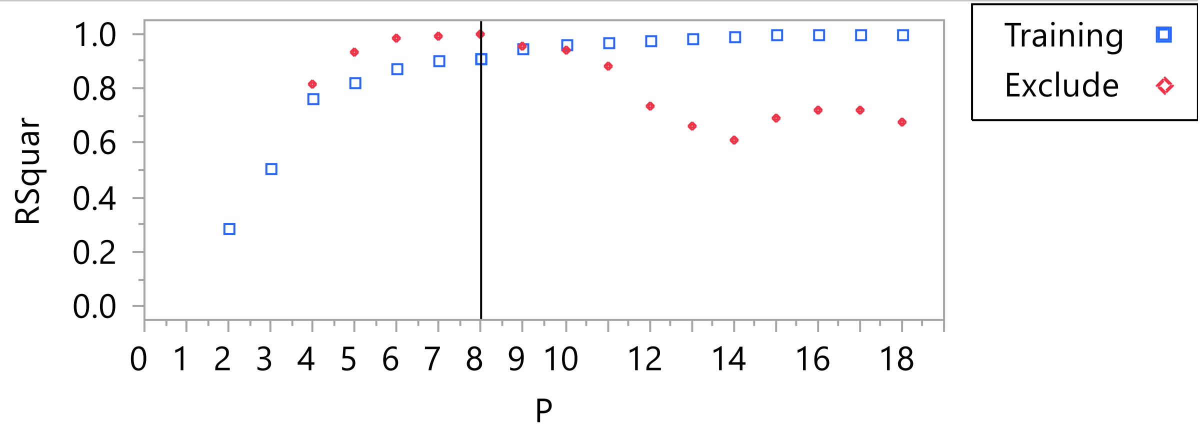


USE MAX VALIDATION R-SQUARE FOR 4 CHECKPOINTS AS STOPPING RULE

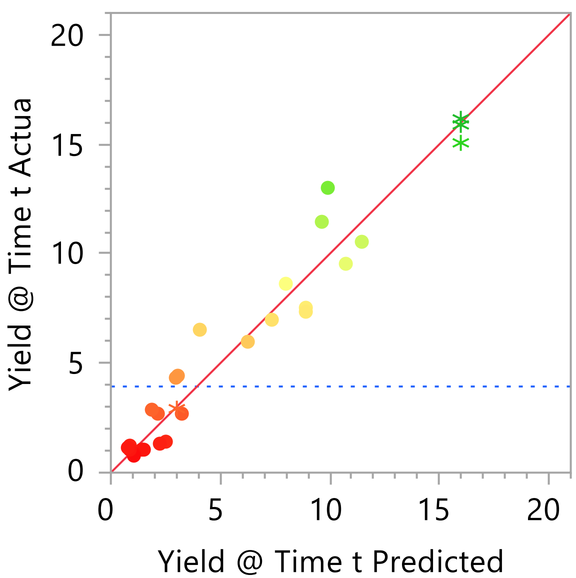
66 TERM QUADRATIC

TRANSFORMED VALUES USED

RSquare History



Actual by Predicted Plot

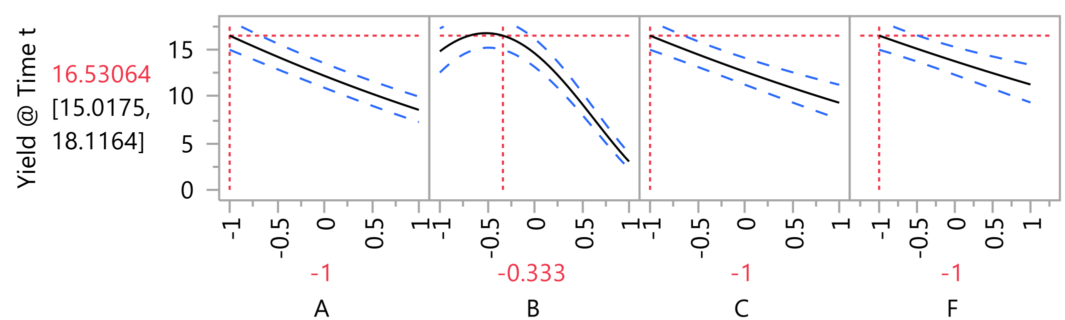


P<.0001 RSq=0.91
RMSE=0.3209

Sorted Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
A	-0.505343	0.072679	-6.95	<.0001 *
B*B	-1.218717	0.17612	-6.92	<.0001 *
B	-0.491041	0.072679	-6.76	<.0001 *
C	-0.231007	0.072679	-3.18	0.0058 *
A*B	0.2306449	0.076075	3.03	0.0079 *
B*C	0.1585526	0.076075	2.08	0.0535
F	-0.110087	0.072679	-1.51	0.1494

Prediction Profiler

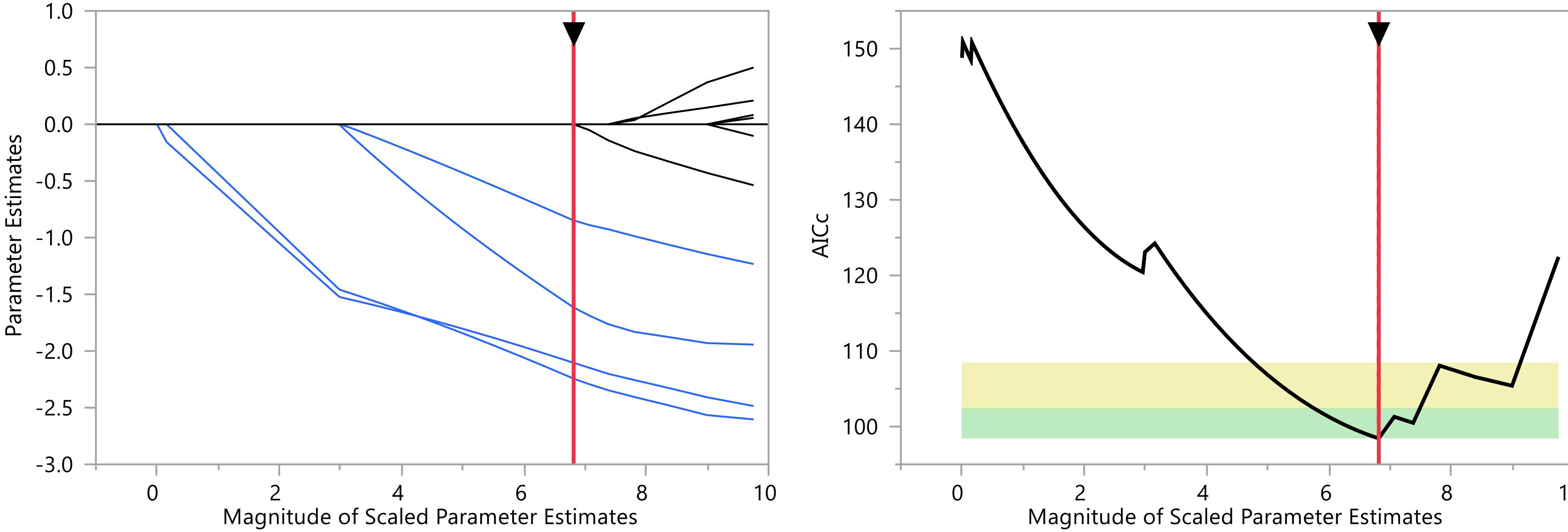


USE AIC CRITERION AS STOPPING RULE

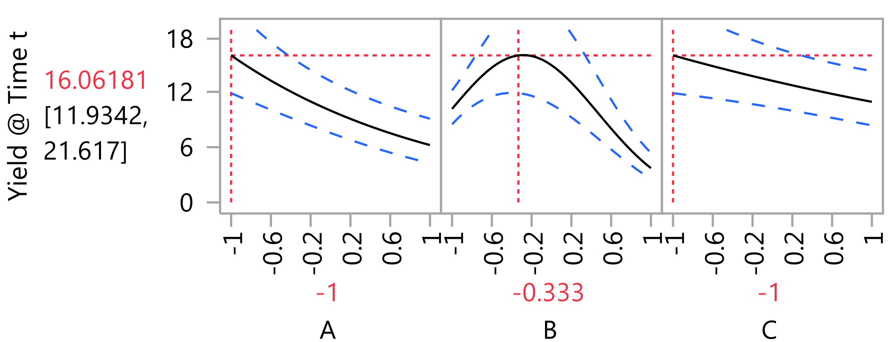
66 TERM QUADRATIC

POISSON DISTRIBUTION USED WITH GENERALIZED REGRESSION

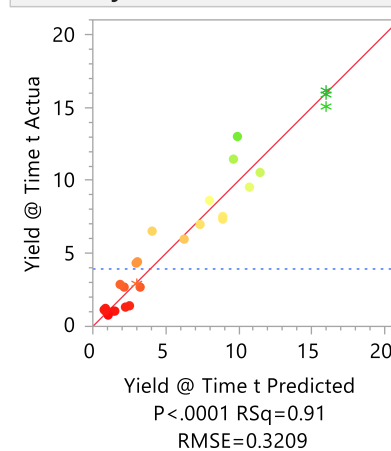
Solution Path



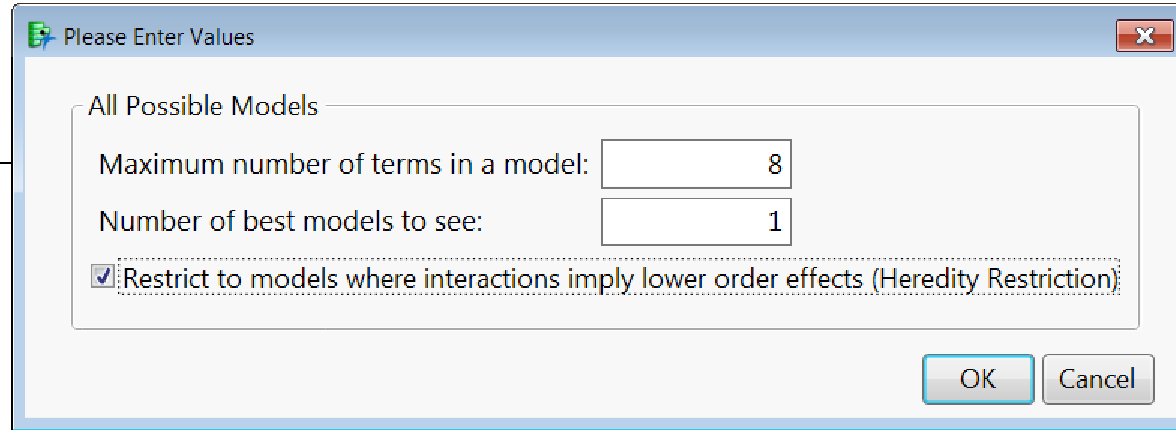
Prediction Profiler

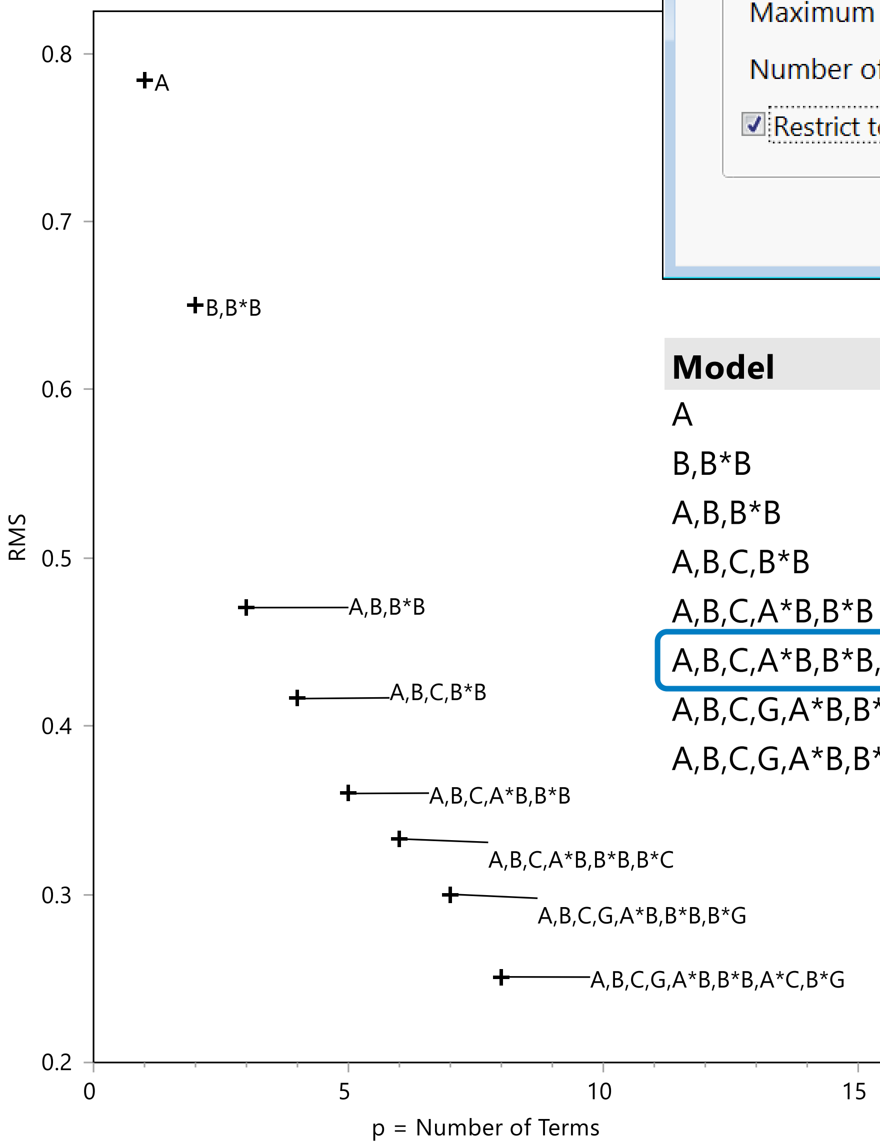


Actual by Predicted Plot



FIT ALL POSSIBLE MODELS UP TO 8 TERMS





Model	Number	RSquare	RMSE	AICc	BIC	
A	1	0.2861	0.7836	61.5162	63.8504	○
B,B*B	2	0.5307	0.6503	54.3550	56.9620	○
A,B,B*B	3	0.7661	0.4704	40.8699	43.4268	○
A,B,C,B*B	4	0.8260	0.4163	37.3830	39.5102	○
A,B,C,A*B,B*B	5	0.8768	0.3599	33.1552	34.4016	○
A,B,C,A*B,B*B,B*C	6	0.9004	0.3329	32.6422	32.4667	◉
A,B,C,G,A*B,B*B,B*G	7	0.9239	0.3000	31.4479	29.1933	○
A,B,C,G,A*B,B*B,A*C,B*	8	0.9501	0.2509	27.3801	22.2375	○

Factors A & B Rsquare = 0.77
Factors A, B & C Rsquare = 0.90
Factors A, B, C & G Rsquare = 0.95

WISDOM FROM BOB

Although your model can fit the data,
it may NOT fit the process from which the data come!

How do I know if my model fits?
" is right?
" adequate?
" accurate?

For me, nothing beats checkpoints!
Do they fall within prediction limits?
What does a plot of actual vs. prediction look like?

Continue to check model predictions over time.
tools wear
seasons change
suppliers and operators change

ALL ANALYSES RANK FACTORS A, B & C AS TOP 3

FACTOR F APPEARS TO BE MOST LIKELY FOURTH FACTOR

- Linear terms only – fourth factor is F
- Linear + Squared terms – fourth factor is D
- Stepwise with min AICc stopping rule – fourth factor is F
- Stepwise with max K-Fold R-Square stopping rule – fourth factor is F
- Stepwise with max Validation R-Square as stopping rule – fourth factor is F
- All possible models – fourth factor is G

- When D & F are in same 5-factor (with A, B, & C) stepwise model, D drops out
- When G & F are in same 5-factor (with A, B, & C) stepwise model, G drops out
- When D & G are in same 5-factor (with A, B, & C) stepwise model, both drop out

- There is an important difference between saying, *“Factor F has no effect.”* and, *“Given the amount of data taken an effect for factor F was not detected.”*

- Augmenting design to support 6-factor quadratic model in A, B, C, D, F & G will
  - help resolve the relative contributions of D, F & G
  - increase the power for all – but especially - the squared terms

IF MORE THAN A FEW FACTORS ARE SIGNIFICANT, THEN AUGMENT DESIGN TO SUPPORT 2ND ORDER MODEL

	A	B	C	D	F	G	Block	Yield @ Time t
14	0	0	0	0	0	0	1	7.49
15	1	1	-1	1	-1	1	1	0.98
16	1	1	1	-1	-1	0	1	0.86
17	-1	1	-1	-1	1	1	1	1.25
18	1	-1	1	1	-1	-1	1	1.03
19	1	1	0	-1	1	-1	1	1.07
20	0	0	0	0	0	0	1	7.33
21	1	-1	-1	0	1	-1	1	2.61
22	-1	-1	0	1	-1	1	1	11.39
23	-1	0	1	-1	1	1	1	12.96
24	1	1	-1	1	1	1	1	1.18
25	1	0	1	1	-1	1	2	•
26	1	-1	0	1	1	0	2	•
27	1	-1	-1	1	0	1	2	•
28	1	-1	0	-1	0	-1	2	•
29	1	0	-1	-1	1	0	2	•
30	1	1	0	-1	0	1	2	•
31	1	0	1	0	1	-1	2	•
32	-1	-1	0	0	1	1	2	•
33	0	0	1	1	-1	-1	2	•
34	-1	-1	1	0	0	0	2	•
35	0	1	1	0	1	0	2	•
36	0	1	-1	1	1	-1	2	•

NOTE: First 13 rows of original design are not shown.

These 12 trials added onto original 24 trials to support full quadratic model in 6 most important factors plus a block effect between original and augmented trials

POWER FOR SQUARED TERMS IN 2ND ORDER MODEL IS INCREASED TO NEAR THAT OF 6-FACTOR RSM DESIGNS

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.273
Block	1	0.983
A	1	0.965
B	-1	0.966
C	1	0.976
D	-1	0.969
F	1	0.975
G	-1	0.961
A*B	1	0.887
A*C	-1	0.881
A*D	1	0.825
A*F	-1	0.915
A*G	1	0.732
B*C	-1	0.728
B*D	1	0.853
B*F	-1	0.859
B*G	1	0.724
C*D	-1	0.872
C*F	1	0.838
C*G	-1	0.778
D*F	1	0.847
D*G	-1	0.838
F*G	1	0.86
A*A	1	0.299
B*B	-1	0.361
C*C	1	0.362
D*D	-1	0.309
F*F	1	0.384
G*G	-1	0.347

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.364
A	1	0.998
B	-1	0.998
C	1	0.998
D	-1	0.998
F	1	0.998
G	-1	0.998
A*A	1	0.527
B*B	-1	0.599
C*C	1	0.582
D*D	-1	0.541
F*F	1	0.573
G*G	-1	0.568

	A	B	C	D	F	G	Block	Yield @ Time t
14	0	0	0	0	0	0	1	7.49
15	1	1	-1	1	-1	1	1	0.98
16	1	1	1	-1	-1	0	1	0.86
17	-1	1	-1	-1	1	1	1	1.25
18	1	-1	1	1	-1	-1	1	1.03
19	1	1	0	-1	1	-1	1	1.07
20	0	0	0	0	0	0	1	7.33
21	1	-1	-1	0	1	-1	1	2.61
22	-1	-1	0	1	-1	1	1	11.39
23	-1	0	1	-1	1	1	1	12.96
24	1	1	-1	1	1	1	1	1.18
25	1	0	1	1	-1	1	2	•
26	1	-1	0	1	1	0	2	•
27	1	-1	-1	1	0	1	2	•
28	1	-1	0	-1	0	-1	2	•
29	1	0	-1	-1	1	0	2	•
30	1	1	0	-1	0	1	2	•
31	1	0	1	0	1	-1	2	•
32	-1	-1	0	0	1	1	2	•
33	0	0	1	1	-1	-1	2	•
34	-1	-1	1	0	0	0	2	•
35	0	1	1	0	1	0	2	•
36	0	1	-1	1	1	-1	2	•

COMPARE AUGMENTED DESIGNS

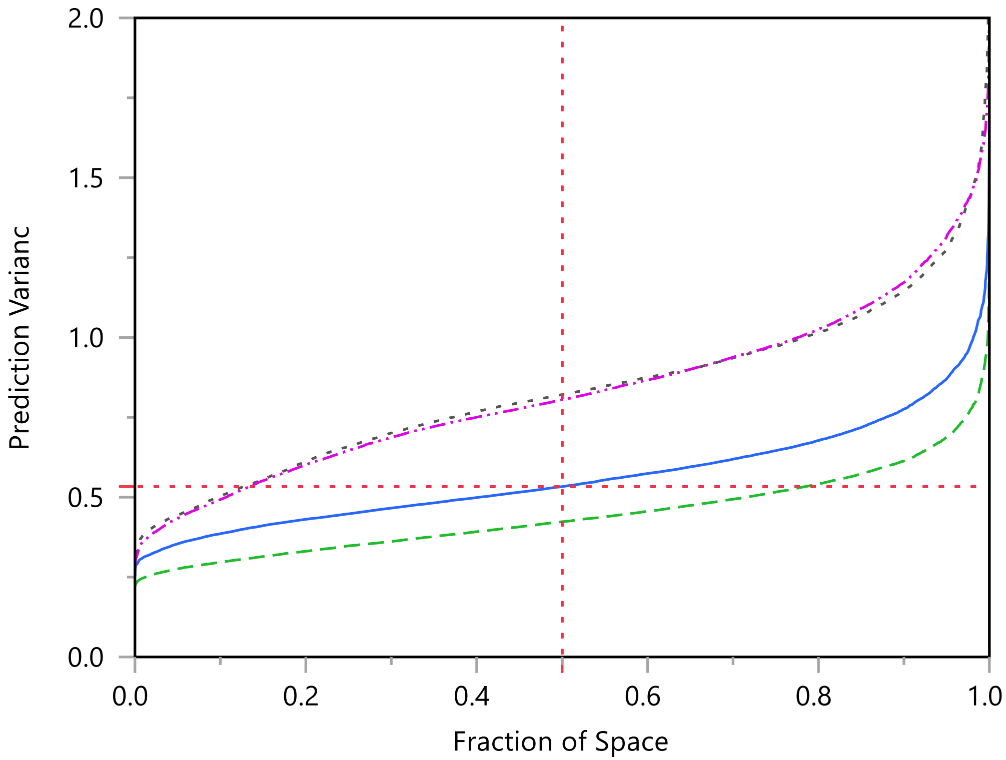
TOP: 10-FACTOR FRACTIONAL FACTORIAL + C.P. AUGMENTED TO SUPPORT FULL QUADRATIC MODEL IN 6 FACTORS
33 + 9 = 42 TOTAL TRIALS

UPPER MIDDLE: 10-FACTOR PLACKET-BURMAN + C.P. AUGMENTED TO SUPPORT FULL QUADRATIC MODEL IN 6 FACTORS
25 + 11 = 36 TOTAL TRIALS

LOWER MIDDLE: 10-FACTOR DEFINITIVE SCREENING AUGMENTED TO SUPPORT FULL QUADRATIC MODEL IN 6 FACTORS
21 + 15 = 36 TOTAL TRIALS

BOTTOM: 6-FACTOR CUSTOM DOE FOR FULL RSM MODEL
34 TOTAL TRIALS

Fraction of Design Space Plot



Design Diagnostics

I Optimal Design	
D Efficiency	40.729
G Efficiency	56.09719
A Efficiency	12.41717
Average Variance of Prediction	0.82307
Design Creation Time (seconds)	0.05

Design Diagnostics

I Optimal Design	
D Efficiency	38.46605
G Efficiency	54.33992
A Efficiency	14.61968
Average Variance of Prediction	0.833744
Design Creation Time (seconds)	0.05

Design Diagnostics

I Optimal Design	
D Efficiency	42.15506
G Efficiency	69.61262
A Efficiency	22.27027
Average Variance of Prediction	0.563765
Design Creation Time (seconds)	0.066667

Design Diagnostics

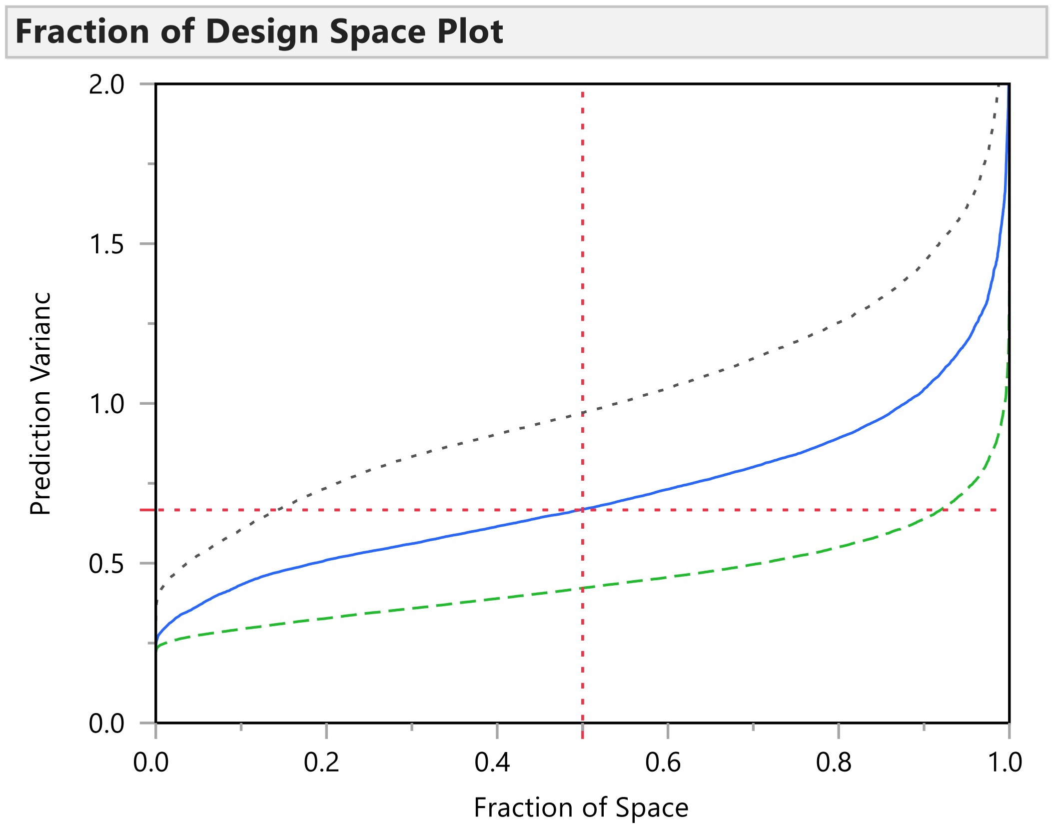
I Optimal Design	
D Efficiency	42.94028
G Efficiency	75.52931
A Efficiency	27.20305
Average Variance of Prediction	0.44424
Design Creation Time (seconds)	0.066667

COMPARE AUGMENTED DESIGNS

TOP: 14-FACTOR FRACTIONAL FACTORIAL + C.P. AUGMENTED TO SUPPORT FULL QUADRATIC MODEL IN 7 FACTORS
33 + 13 = 46 TOTAL TRIALS

MIDDLE: 14-FACTOR DEFINITIVE SCREENING AUGMENTED TO SUPPORT FULL QUADRATIC MODEL IN 7 FACTORS
29 + 17 = 46 TOTAL TRIALS

BOTTOM: 7-FACTOR CUSTOM DOE FOR FULL RSM MODEL
42 TOTAL TRIALS



Design Diagnostics

I Optimal Design	
D Efficiency	37.352
G Efficiency	48.68453
A Efficiency	11.13939
Average Variance of Prediction	1.006709
Design Creation Time (seconds)	0.133333

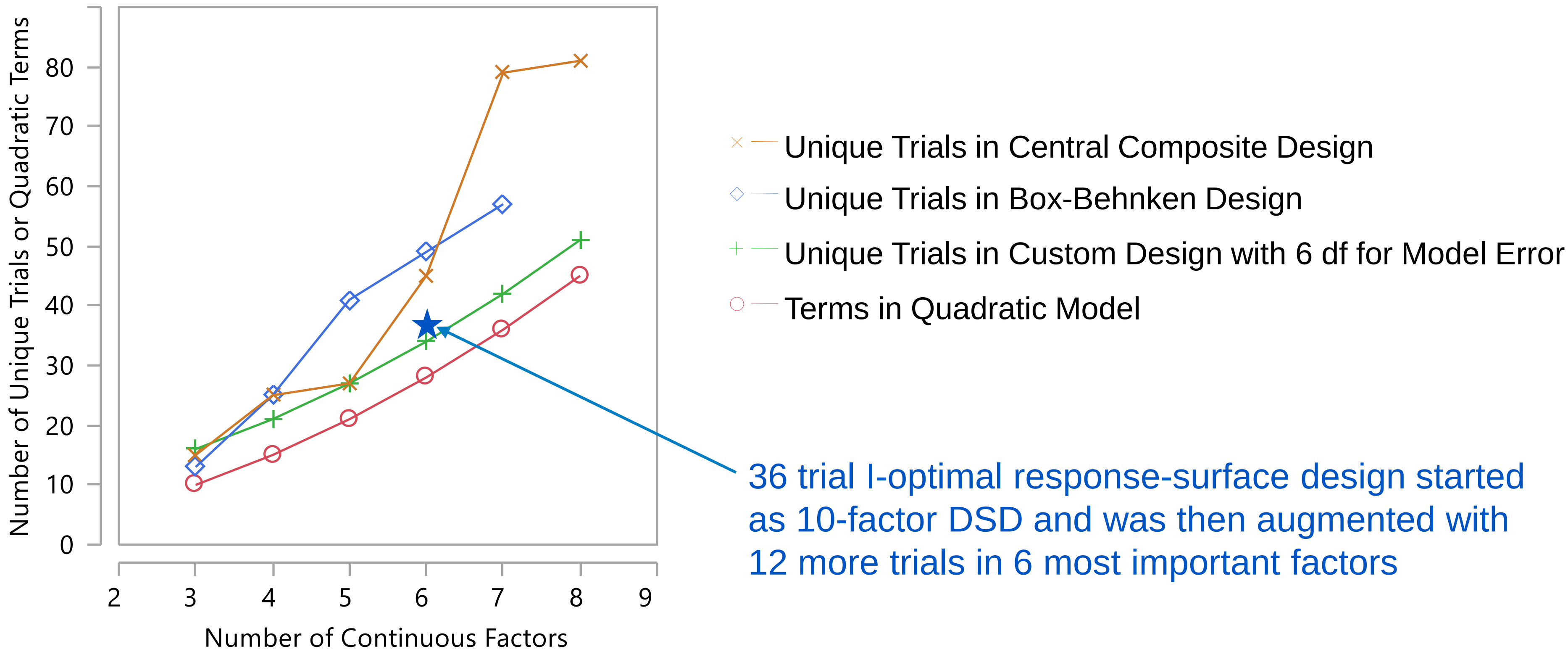
Design Diagnostics

I Optimal Design	
D Efficiency	36.69963
G Efficiency	58.39688
A Efficiency	15.61337
Average Variance of Prediction	0.714178
Design Creation Time (seconds)	0.133333

Design Diagnostics

I Optimal Design	
D Efficiency	41.03495
G Efficiency	71.04153
A Efficiency	27.70772
Average Variance of Prediction	0.449918
Design Creation Time (seconds)	0.216667

NUMBER OF UNIQUE TRIALS FOR 3 RESPONSE-SURFACE DESIGNS AND NUMBER OF QUADRATIC MODEL TERMS VS. NUMBER OF CONTINUOUS FACTORS



If generally running 3, 4 or 5-factor fractional-factorial designs…

1. How many interactions are you not investigating?
2. How many more trials needed to fit curvature?
3. Consider two stages: Definitive Screening + Augmentation

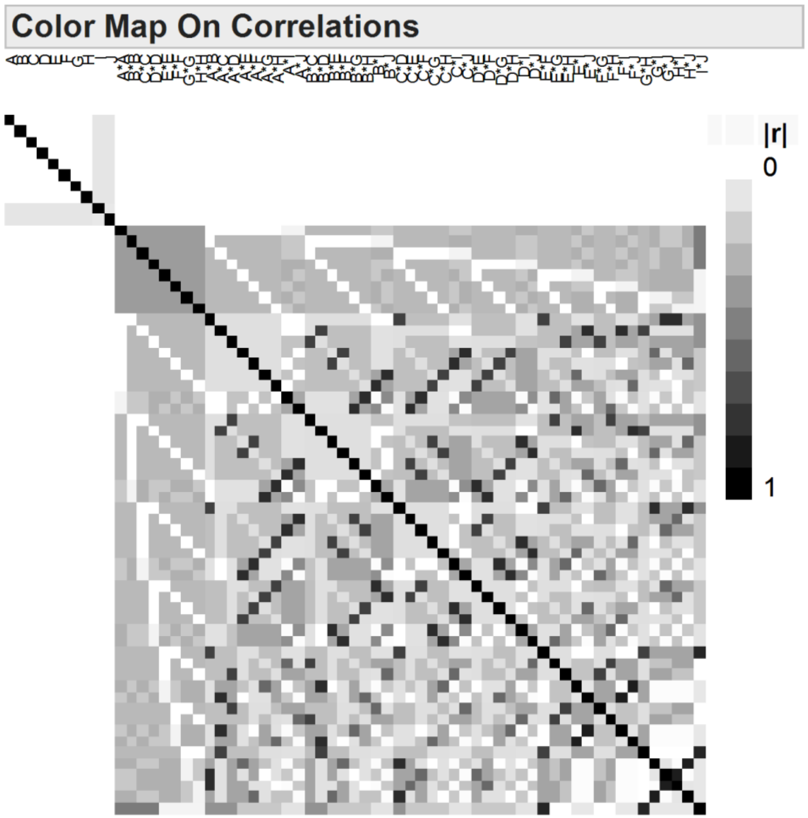
- ***Definitive Screening Designs***
  - Efficiently estimate main and quadratic effects for no more and **often fewer trials than traditional designs**
  - If only a few factors are important the design may collapse into a **"one-shot"** design that supports a response-surface model
  - If many factors are important the design can be **augmented** to support a response-surface model
  - Case study for a **10-variable process** shows that it can be **optimized in just 23 unique trials**

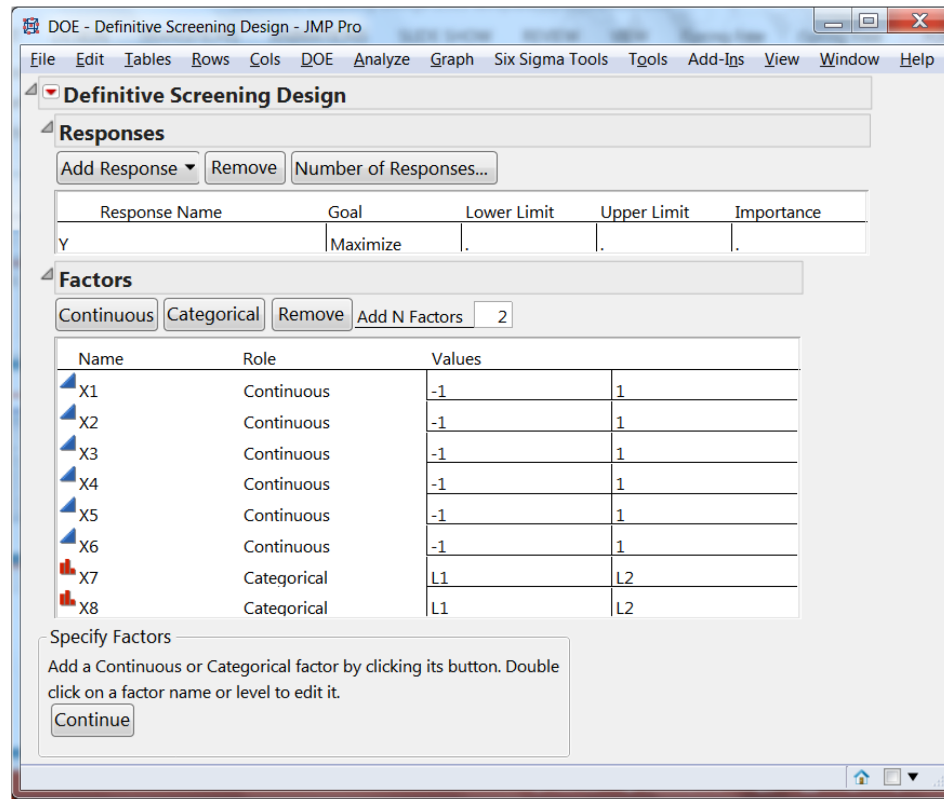
Thanks.
Questions or comments?

TOM.DONNELLY@JMP.COM

JMP 11 DEFINITIVE SCREENING DESIGN COLOR MAPS FOR 8-CONTINUOUS, 2-CATEGORICAL FACTOR

De-alias 2-f Interactions and Categorical Factors



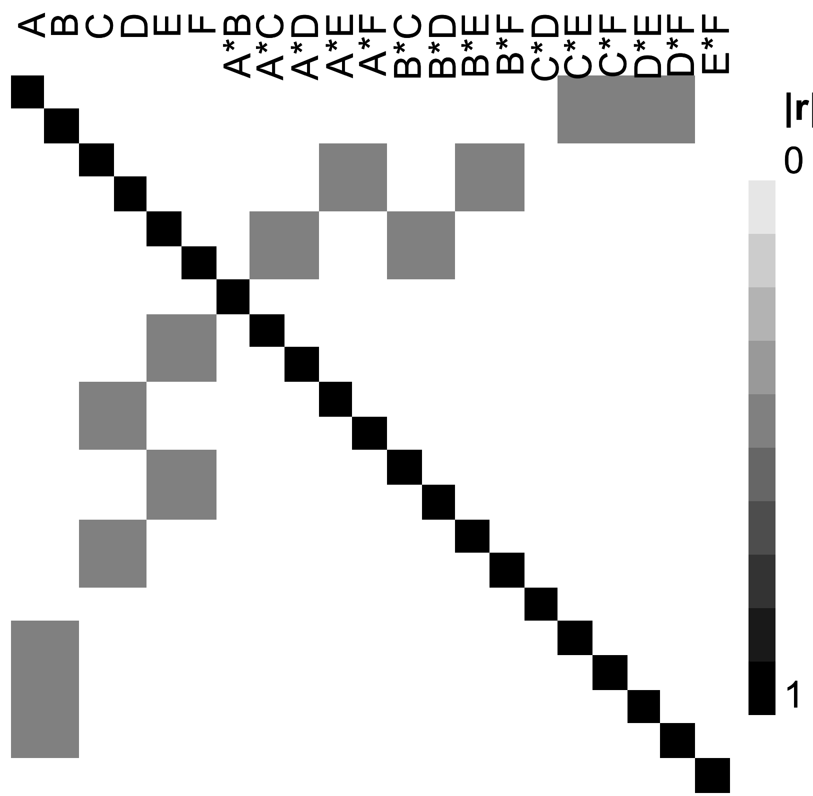


6-FACTOR, 16-TRIAL, NON-REGULAR FRACTIONAL FACTORIAL ("NO CONFOUNDING" DESIGN)

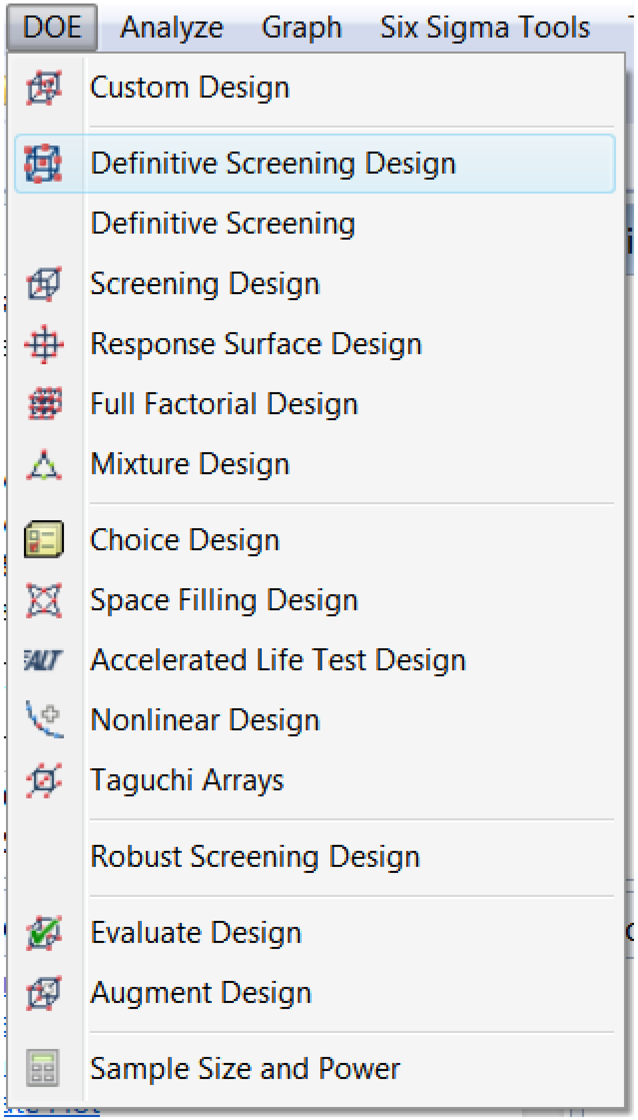
Jones, B. and Montgomery, D., (2010) "Alternatives to Resolution IV Screening Designs in 16 Runs." *International Journal of Experimental Design and Process Optimization*, 2010; Vol. 1 No. 4: 285-295.

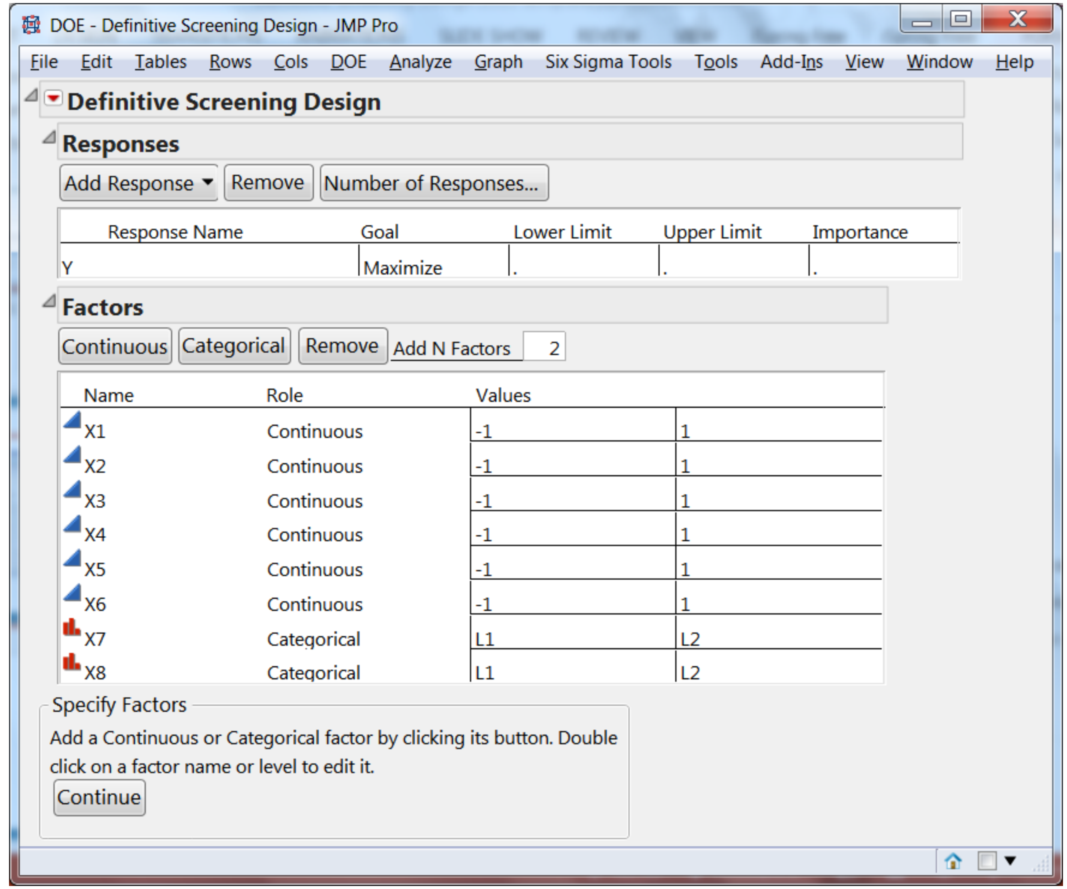
	A	B	C	D	E	F
1	1	1	1	1	1	1
2	1	1	-1	-1	-1	-1
3	-1	-1	1	1	-1	-1
4	-1	-1	-1	-1	1	1
5	1	1	1	-1	1	-1
6	1	1	-1	1	-1	1
7	-1	-1	1	-1	-1	1
8	-1	-1	-1	1	1	-1
9	1	-1	1	1	1	-1
10	1	-1	-1	-1	-1	1
11	-1	1	1	1	-1	1
12	-1	1	-1	-1	1	-1
13	1	-1	1	-1	-1	-1
14	1	-1	-1	1	1	1
15	-1	1	1	-1	1	1
16	-1	1	-1	1	-1	-1

Color Map On Correlations



WITH JMP 11 USE DEFINITIVE SCREENING ON DOE MENU



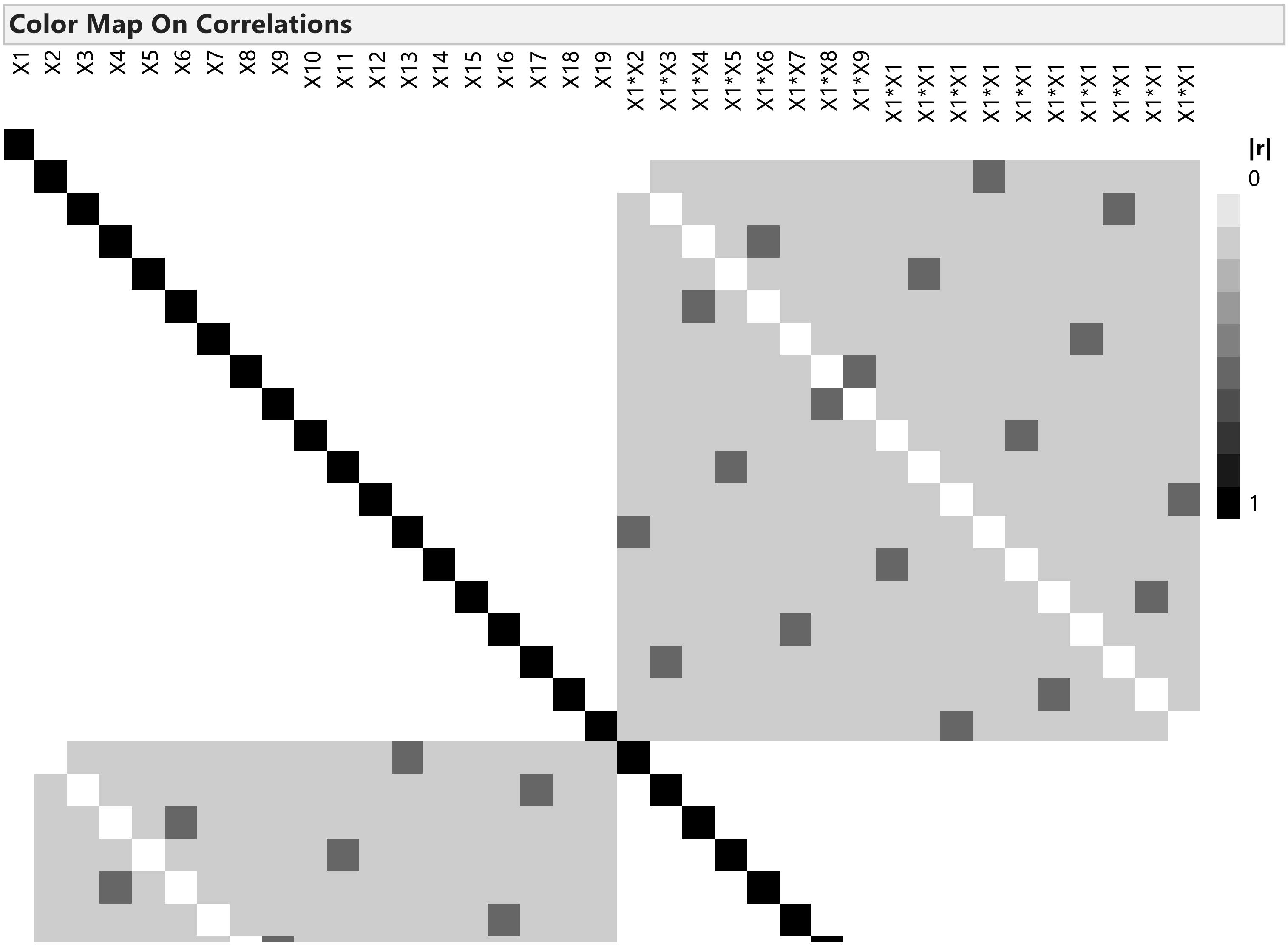


ANALYSIS STRATEGIES

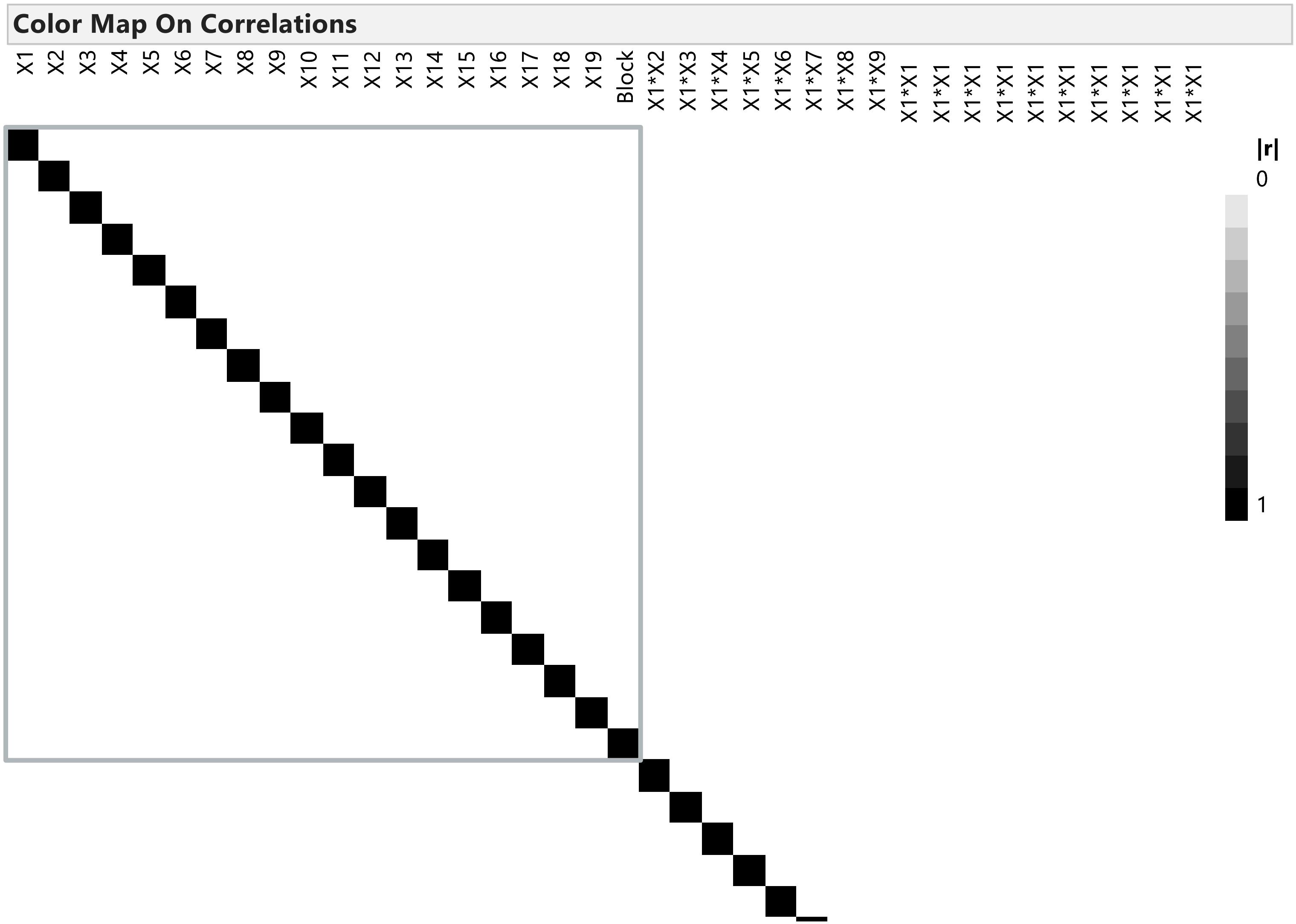
- **Visual Tools:**
  - **Distribution** – click on “good” and “bad” response values to see correlations with factor settings
  - **Graph Builder** – Y vs. X graphs – all data, summarized data, fit line, smoother
    - Drop factors side by side or alternatively (for coded factors) stack factors then replot
    - Use Overlay field to look at possible interactions between two factors

- **Analytic Tools:**
  - **Conservative**: Main Effects fit – look at Scaled estimates
    - Consider adding interactions among significant factors using Effects Heredity and Sparsity
  - **Aggressive**: Strepwise with various stopping criteria
    - AICc, BIC, K-fold, Excluded checkpoints,
    - Fit All Possible Models
  - **Analytic Output:**
    - Stepwise Histories – Criterion or Rsquare
    - Actual vs. Predicted with Graph Builder – Col Switch different models
    - Create All Possible Models Table – Plot four metrics using Overlap Plot

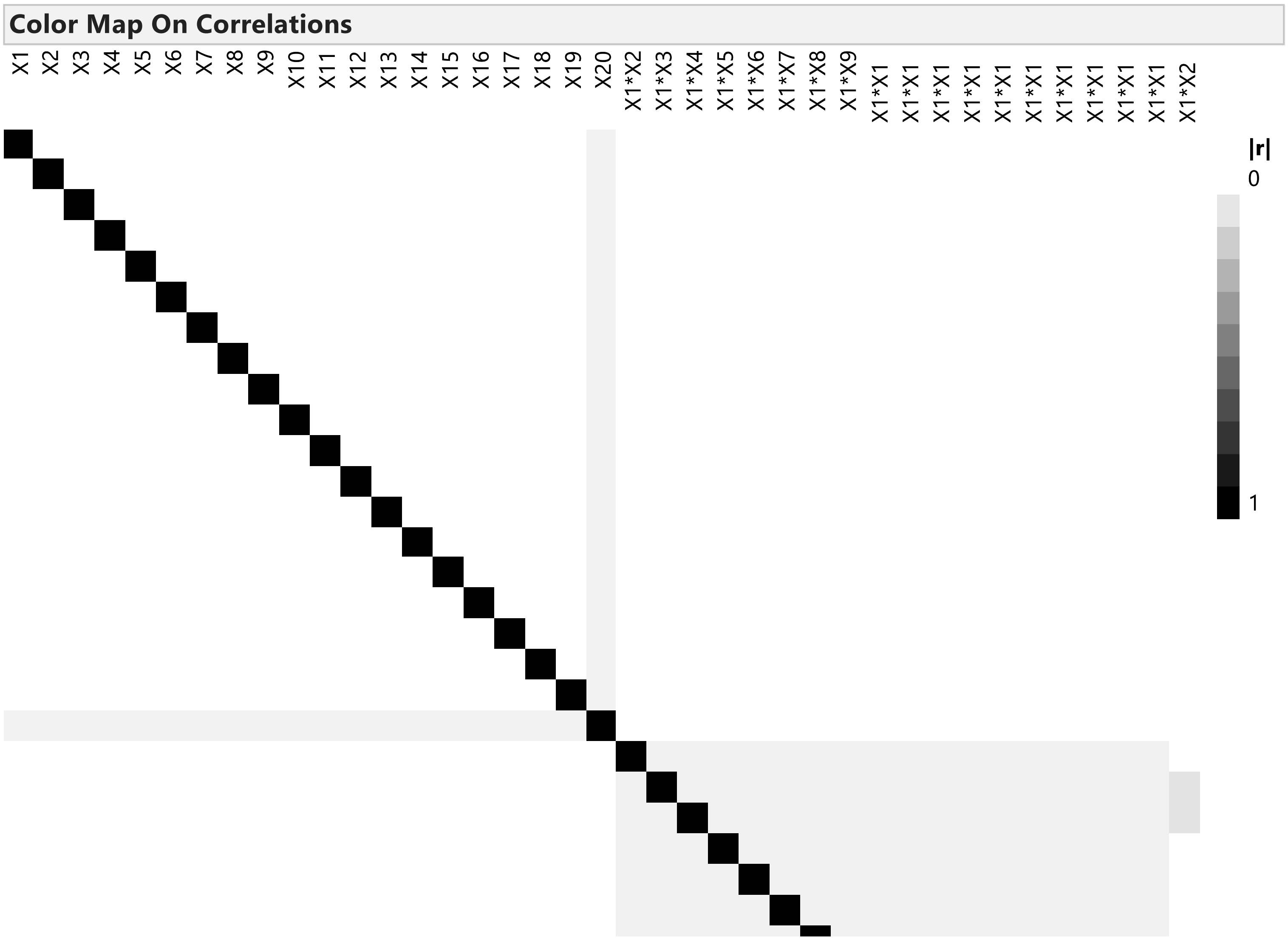
COLOR MAP FOR 20-TRIAL PLACKETT-BURMAN DESIGN WITH 19 CONTINUOUS FACTORS



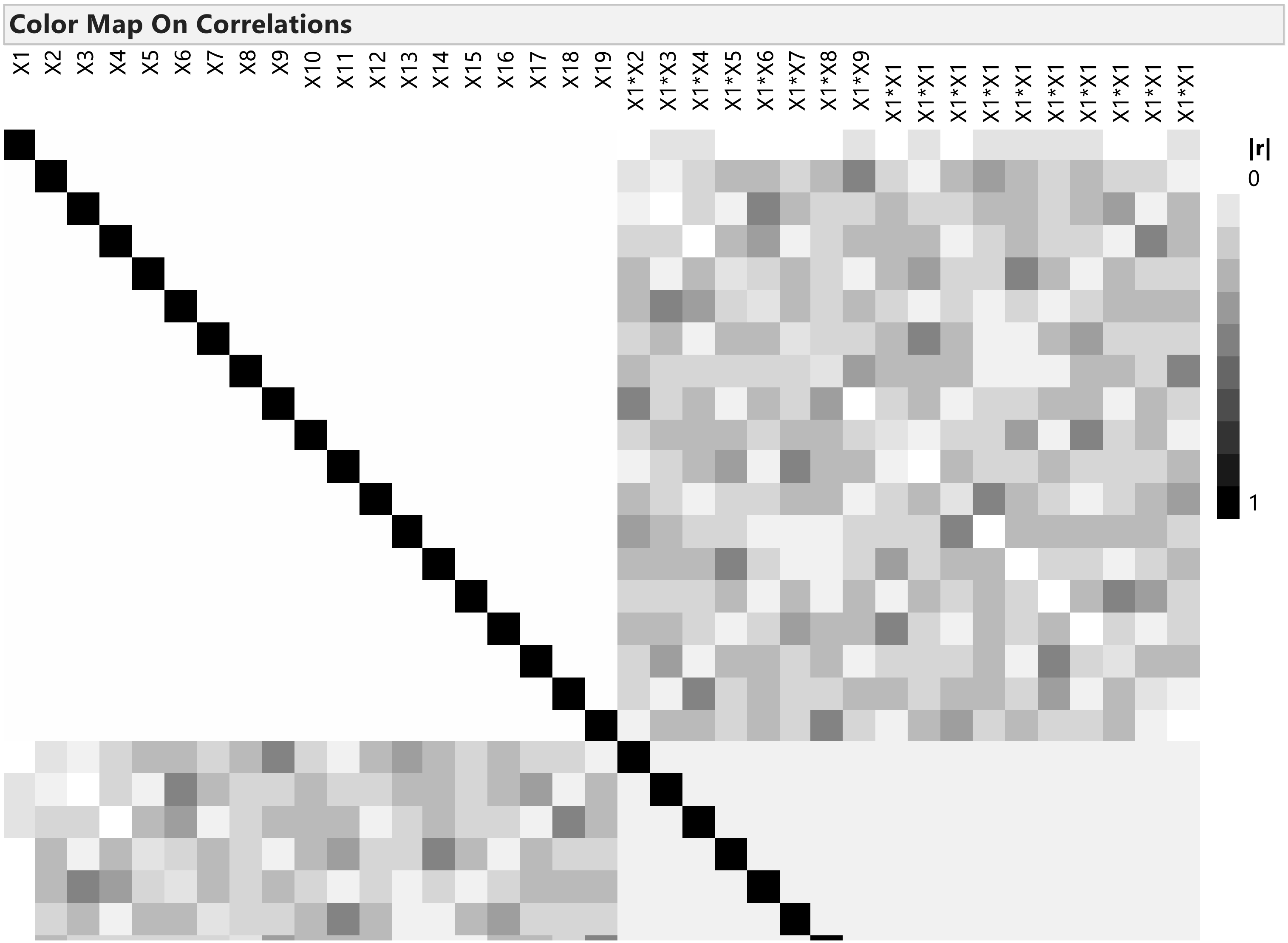
COLOR MAP FOR 40-TRIAL FOLD-OVER PLACKETT-BURMAN DESIGN WITH 19 CONTINUOUS FACTORS AND 20TH BLOCK FACTOR



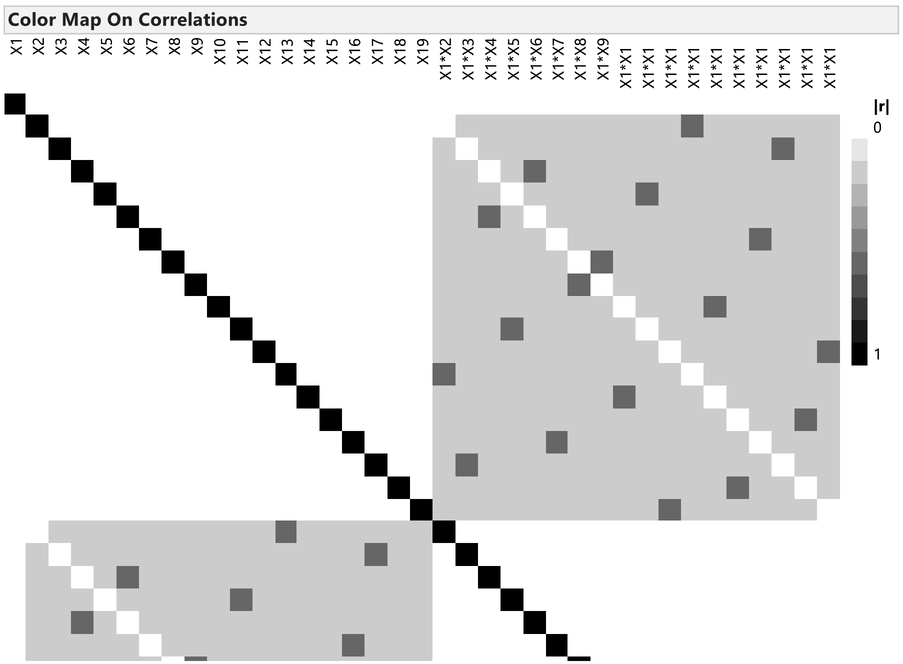
COLOR MAP FOR A 42-TRIAL DEFINITIVE SCREENING DESIGN WITH 19 CONTINUOUS FACTORS AND 1 TWO-LEVEL CATEGORICAL FACTOR

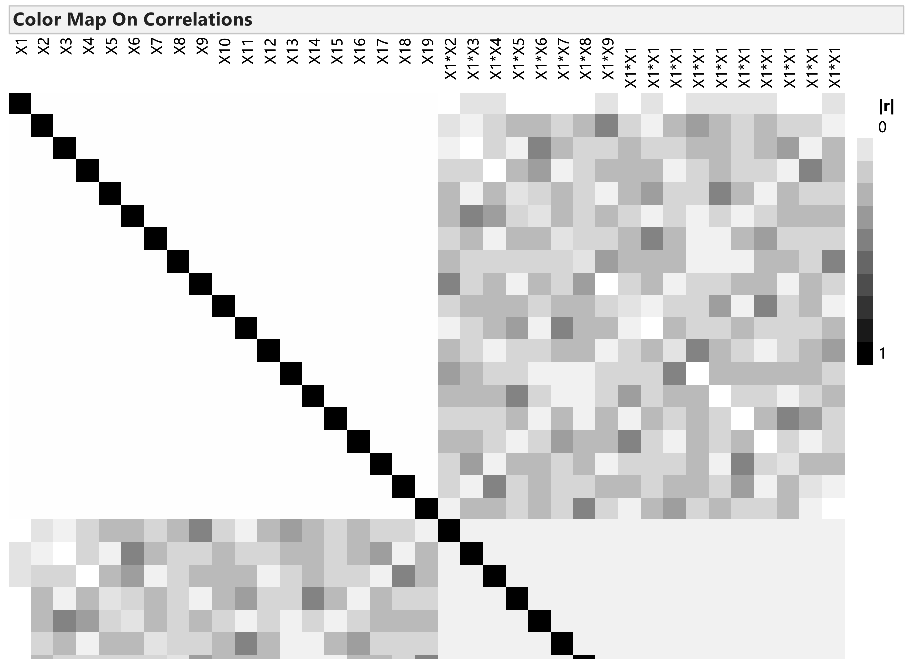


COLOR MAP FOR 21-TRIAL HALF OF 42-TRIAL DSD WITH 19 CONTINUOUS FACTORS SPLIT ON 20TH CATEGORICAL FACTOR

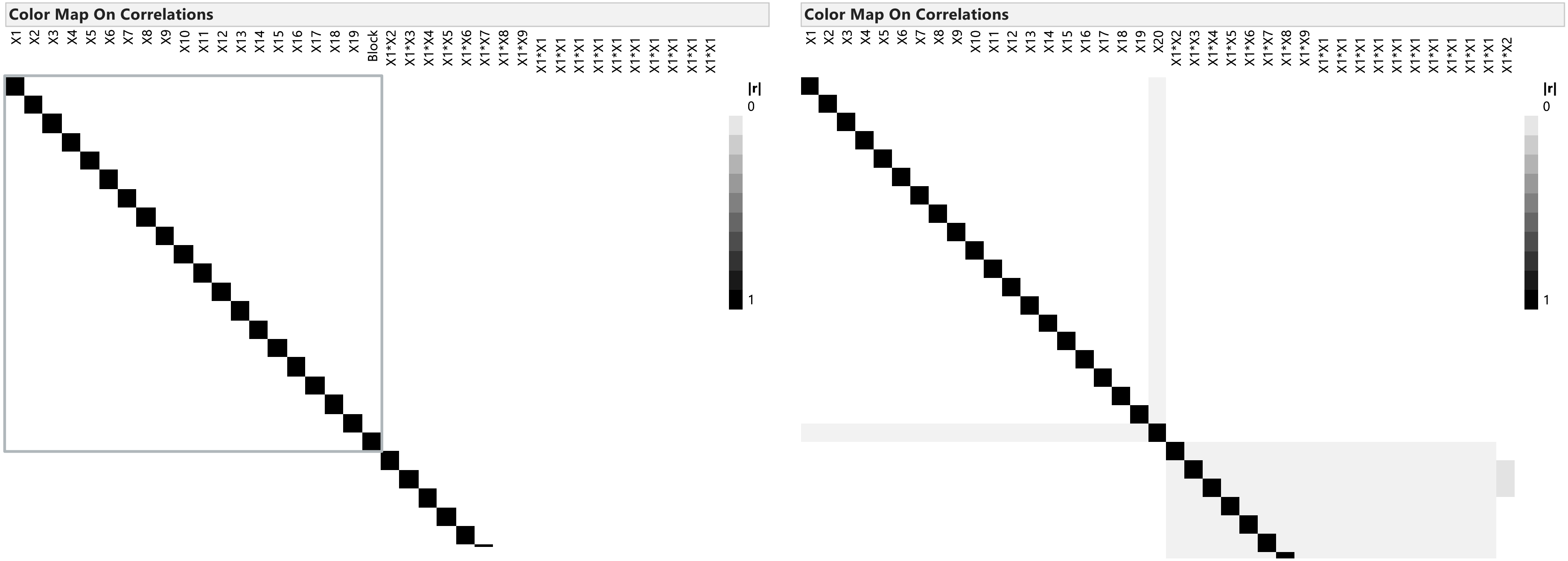


COLOR MAP FOR 20-TRIAL PLACKETT-BURMAN DESIGN (LEFT) AND 21-TRIAL HALF OF 42-TRIAL DSD (RIGHT) BOTH WITH 19 CONTINUOUS FACTORS





COLOR MAP FOR A 40-TRIAL FOLD-OVER PLACKET-BURMAN DESIGN (LEFT) AND A 42-TRIAL DEFINITIVE SCREENING DESIGN (RIGHT) WITH 19 CONTINUOUS AND 1 TWO-LEVEL BLOCK/CATEGORICAL FACTOR



For designs containing only continuous factors, compare these properties of definitive screening designs versus standard screening designs:

- Main effects are orthogonal to two-factor interactions.
  - Definitive Screening Designs: Always
  - Standard Screening Designs: Only for Resolution IV or higher

- No two-factor interaction is completely confounded with any other two-factor interaction.
  - Definitive Screening Designs: Always
  - Standard Screening Designs: Only for Resolution V or higher

- All quadratic effects* are estimable in models containing only main and quadratic effects.
  - Definitive Screening Designs: Always
  - Standard Screening Designs: Never

* When quadratic effects are mentioned, the standard screening designs are assumed to have center points.